

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 11:11:40 ; Search time 80.4247 Seconds
(without alignments)
7140.092 Million cell updates/sec

Title: US-09-845-416-10_COPY_1800_3100
Perfect score: 1301
Sequence: 1 cgacttcagcagcttcaga.....ccgagctgctccaattg 1301

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 segs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|--|
| 1 | 1301 | 100.0 | 5952 | 4 | US-09-687-875A-1 Sequence 1, Appli |
| 2 | 1290 | 99.2 | 13977 | 4 | US-09-484-970B-60 Sequence 60, Appli |
| 3 | 1135.2 | 87.3 | 19307 | 3 | US-08-836-022A-10 Sequence 10, Appli |
| 4 | 1135.2 | 87.3 | 19307 | 3 | US-09-427-048A-10 Sequence 10, Appli |
| 5 | 553.8 | 42.6 | 6045 | 4 | US-09-091-501B-7 Sequence 7, Appli |
| 6 | 553.8 | 42.6 | 10320 | 4 | US-09-091-501B-9 Sequence 9, Appli |
| 7 | 50 | 3.8 | 7218 | 1 | US-08-232-463-14 Sequence 14, Appli |
| 8 | 44.2 | 3.4 | 2574 | 4 | US-09-668-313A-10 Sequence 10, Appli |
| 9 | 42.8 | 3.3 | 1690 | 4 | US-09-620-312D-69 Sequence 69, Appli |
| 10 | 42.8 | 3.3 | 7812 | 3 | US-09-368-590-1 Sequence 1, Appli |
| 11 | 38.6 | 3.0 | 1995 | 1 | US-08-425-069-3 Sequence 3, Appli |
| 12 | 38.6 | 3.0 | 1995 | 2 | US-08-317-844B-3 Sequence 3, Appli |
| 13 | 38.4 | 3.0 | 7672 | 4 | US-09-220-132-24 Sequence 24, Appli |
| 14 | 38.2 | 2.9 | 428 | 4 | US-09-668-313A-3 Sequence 3, Appli |
| 15 | 38.2 | 2.9 | 4439 | 4 | US-09-668-313A-17 Sequence 17, Appli |
| 16 | 34.6 | 2.7 | 2277 | 1 | US-08-676-967-5 Sequence 5, Appli |
| 17 | 34.6 | 2.7 | 2277 | 1 | US-08-676-974-5 Sequence 5, Appli |
| 18 | 34.6 | 2.7 | 2277 | 2 | US-09-098-487-5 Sequence 5, Appli |
| 19 | 34.4 | 2.6 | 1047 | 4 | US-09-671-950-1 Sequence 1, Appli |
| 20 | 34.4 | 2.6 | 1047 | 4 | US-09-671-950-3 Sequence 3, Appli |
| 21 | 34.4 | 2.6 | 1047 | 4 | US-09-671-950-5 Sequence 5, Appli |
| 22 | 34.4 | 2.6 | 1047 | 4 | US-09-671-950-7 Sequence 7, Appli |
| 23 | 34.4 | 2.6 | 1047 | 4 | US-09-671-950-9 Sequence 9, Appli |
| 24 | 34.4 | 2.6 | 1047 | 4 | US-09-671-950-11 Sequence 11, Appli |
| 25 | 34.4 | 2.6 | 1047 | 4 | US-09-671-950-13 Sequence 13, Appli |
| 26 | 34.2 | 2.6 | 750 | 4 | US-08-961-527-370 Sequence 370, Appli |
| 27 | 34.2 | 2.6 | 1620 | 4 | US-08-858-207A-56 Sequence 56, Appli |

| | | | | | |
|----|------|-----|-------|---|---|
| 28 | 34.2 | 2.6 | 2800 | 2 | US-08-874-138-1 Sequence 1, Appli |
| 29 | 34.2 | 2.6 | 2800 | 2 | US-08-874-138-5 Sequence 5, Appli |
| 30 | 34.2 | 2.6 | 2800 | 3 | US-08-879-941-1 Sequence 1, Appli |
| 31 | 34.2 | 2.6 | 2800 | 3 | US-08-879-941-3 Sequence 3, Appli |
| 32 | 34.2 | 2.6 | 2800 | 4 | US-09-747-116-1 Sequence 1, Appli |
| 33 | 34.2 | 2.6 | 2800 | 4 | US-09-747-116-3 Sequence 3, Appli |
| 34 | 34.2 | 2.6 | 32768 | 4 | US-08-961-527-71 Sequence 71, Appli |
| 35 | 34 | 2.6 | 648 | 4 | US-09-252-991A-10033 Sequence 10033, A |
| 36 | 34 | 2.6 | 762 | 4 | US-09-252-991A-9821 Sequence 9821, Ap |
| 37 | 34 | 2.6 | 1644 | 4 | US-09-252-991A-10161 Sequence 10161, A |
| 38 | 34 | 2.6 | 2235 | 3 | US-09-153-804-2 Sequence 2, Appli |
| 39 | 33.8 | 2.6 | 1603 | 1 | US-08-625-209A-1 Sequence 1, Appli |
| 40 | 33.8 | 2.6 | 3489 | 2 | US-08-728-323A-1 Sequence 2, Appli |
| 41 | 33.8 | 2.6 | 3489 | 4 | US-09-298-568-1 Sequence 1, Appli |
| 42 | 33.8 | 2.6 | 3489 | 4 | US-09-410-399-1 Sequence 1, Appli |
| 43 | 33.8 | 2.6 | 32207 | 2 | US-08-770-379-20 Sequence 20, Appli |
| 44 | 33.8 | 2.6 | 32207 | 3 | US-08-757-669A-20 Sequence 20, Appli |
| 45 | 33.8 | 2.6 | 32207 | 4 | US-09-230-371A-20 Sequence 20, Appli |

ALIGNMENTS

| | | | | | | | | | |
|---|------|--|------|--|--|--|--|--|--|
| RESULT 1 | | | | | | | | | |
| US-09-687-875A-1 | | | | | | | | | |
| Sequence 1, Application US/09687875A | | | | | | | | | |
| Patent No. 6544786 | | | | | | | | | |
| GENERAL INFORMATION: | | | | | | | | | |
| APPLICANT: Xiao, Xiao | | | | | | | | | |
| APPLICANT: Liu, Paul | | | | | | | | | |
| TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPICED | | | | | | | | | |
| FILE REFERENCE: 00792 | | | | | | | | | |
| CURRENT APPLICATION NUMBER: US/09/687, 875A | | | | | | | | | |
| CURRENT FILING DATE: 2000-10-13 | | | | | | | | | |
| PRIOR APPLICATION NUMBER: 60/158, 868 | | | | | | | | | |
| PRIOR FILING DATE: 1999-10-15 | | | | | | | | | |
| NUMBER OF SEQ ID NOS: 22 | | | | | | | | | |
| SOFTWARE: PatentIn version 3.1 | | | | | | | | | |
| SEQ ID NO 1 | | | | | | | | | |
| LENGTH: 5952 | | | | | | | | | |
| TYPE: DNA | | | | | | | | | |
| ORGANISM: Homo sapiens | | | | | | | | | |
| FEATURE: | | | | | | | | | |
| NAME/KEY: misc feature | | | | | | | | | |
| LOCATION: (2897)..(2898) | | | | | | | | | |
| OTHER INFORMATION: S4 junction site | | | | | | | | | |
| NAME/KEY: misc feature | | | | | | | | | |
| LOCATION: (3198)..(3199) | | | | | | | | | |
| OTHER INFORMATION: S2 junction site | | | | | | | | | |
| US-09-687-875A-1 | | | | | | | | | |
| Query Match | | | | | | | | | |
| Best Local Similarity 100.0%; Score 1301; DB 4; Length 5952; | | | | | | | | | |
| Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | | | | | | | |
| QY | 1 | CGACTTCCAGCAGTTTCAGAAAGCAGACGATGTACATATAGGGCCCTTCAAGAGGAATTGAA | 60 | | | | | | |
| DB | 3402 | CGACTTCCAGCAGTTTCAGAAAGCAGACGATGTACATATAGGGCCCTTCAAGAGGAATTGAA | 3461 | | | | | | |
| QY | 61 | AACTAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA | 120 | | | | | | |
| DB | 3462 | AACTAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA | 3521 | | | | | | |
| QY | 121 | GCCTTGAAGAGACTAGAAACTCTACCCAGAGCCAGAGAGCTGCCTCTGAGAGAG | 180 | | | | | | |
| DB | 3522 | GCCTTGAAGAGACTAGAAACTCTACCCAGAGCCAGAGAGCTGCCTCTGAGAGAG | 3581 | | | | | | |
| QY | 181 | AGCCCAAGATGTCACTGGCTTCTACGAAGACAGGCTGAGAGGTCAATCTGAGTGGGA | 240 | | | | | | |
| DB | 3582 | AGCCCAAGATGTCACTGGCTTCTACGAAGACAGGCTGAGAGGTCAATCTGAGTGGGA | 3641 | | | | | | |
| QY | 241 | AAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAATAGATGAGACCCCTTGAAGACT | 300 | | | | | | |

Db 3642 AAAATTGAACCTGCACTCCGCTGAAGTGGCAGAGAAAAATAGATGAGACCCTTGAAGAAGT 3701
QY 301 CCAGAACTTCAAGAGGCCAGGATGAGCTGAGCTTCAAGCTGCGCAAGCTGAGGTGAT 360
Db 3702 CCAGAACTTCAAGAGGCCAGGATGAGCTGAGCTTCAAGCTGCGCAAGCTGAGGTGAT 3761
QY 361 CAAGGATCCTGGCAGCCCGTGGGCGATCTCTCATTTGACTTCTCTCAAGATCACTCGA 420
Db 3762 CAAGGATCCTGGCAGCCCGTGGGCGATCTCTCATTTGACTTCTCTCAAGATCACTCGA 3821
QY 421 GAAAGTCAAGGCACCTTGGAGAGAAATTGGCCTCTGAAGAAGAACGTGAGCCAGTCAA 480
Db 3822 GAAAGTCAAGGCACCTTGGAGAGAAATTGGCCTCTGAAGAAGAACGTGAGCCAGTCAA 3881
QY 481 TGACCTTGTCTGCCAGCTTACCATTGGGCAATTCAGCTCTCAACGTATACTCAAGCAC 540
Db 3882 TGACCTTGTCTGCCAGCTTACCATTGGGCAATTCAGCTCTCAACGTATACTCAAGCAC 3941
QY 541 TCTGGAAGACCTGAACACCAAGATGAAAGCTTTCGAGGTGGCCGTGAGGACCGAGTCAG 600
Db 3942 TCTGGAAGACCTGAACACCAAGATGAAAGCTTTCGAGGTGGCCGTGAGGACCGAGTCAG 4001
QY 601 GCAGCTGATGAAGGCCACAGGACCTTGTGTCAGCATCTCAGCACTTCTTTCACAGTC 660
Db 4002 GCAGCTGATGAAGGCCACAGGACCTTGTGTCAGCATCTCAGCACTTCTTTCACAGTC 4061
QY 661 TGTCCAGGTCCTGGAGAGAGCCATCTGCGCAAAAGAGTCCCTACTATATCAACCA 720
Db 4062 TGTCCAGGTCCTGGAGAGAGCCATCTGCGCAAAAGAGTCCCTACTATATCAACCA 4121
QY 721 CGAGACTCAAAACAACCTTGTGGAGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTAC 780
Db 4122 CGAGACTCAAAACAACCTTGTGGAGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTAC 4181
QY 781 TGACCTGAATAATGTCAAGATTCAGCTTATAGACTGCGATGAAGTCCGAGAGACTGCA 840
Db 4182 TGACCTGAATAATGTCAAGATTCAGCTTATAGACTGCGATGAAGTCCGAGAGACTGCA 4241
QY 841 GAAGGCCCTTGTGATCTCTGAGCCTGTGAGCTGCAATGTGATGCTTGAACAGCA 900
Db 4242 GAAGGCCCTTGTGATCTCTGAGCCTGTGAGCTGCAATGTGATGCTTGAACAGCA 4301
QY 901 CAACCTCAAGCAAAATGACAGCCCATGATATCTGCAANTATTAATTGTTGACCAAC 960
Db 4302 CAACCTCAAGCAAAATGACAGCCCATGATATCTGCAANTATTAATTGTTGACCAAC 4361
QY 961 TATTTATGACCCGCTGAGCAAGACACAACAATTTGTTCAACGTCCCTCTCTGGGTGA 1020
Db 4362 TATTTATGACCCGCTGAGCAAGACACAACAATTTGTTCAACGTCCCTCTCTGGGTGA 4421
QY 1021 TATGTGTCTGAAGTGGCTGTGATGTTATGATACCGGACGAACAGGAGAGATCCGTGT 1080
Db 4422 TATGTGTCTGAAGTGGCTGTGATGTTATGATACCGGACGAACAGGAGAGATCCGTGT 4481
QY 1081 CCTGTCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACAATTGGAAGACAAGTACAG 1140
Db 4482 CCTGTCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACAATTGGAAGACAAGTACAG 4541
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGATTTTGTGACCAAGCGAGGCTGGGCT 1200
Db 4542 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGATTTTGTGACCAAGCGAGGCTGGGCT 4601
QY 1201 CCTTCTGATGATTTATCCAAATTCAGAGACAGTTGGTGAAGTGCATCTTTGGGGG 1260
Db 4602 CCTTCTGATGATTTATCCAAATTCAGAGACAGTTGGTGAAGTGCATCTTTGGGGG 4661
QY 1261 CAGTAACATGAGCCAAAGTGTCCGAGCTGCTTCCAATTG 1301
Db 4662 CAGTAACATGAGCCAAAGTGTCCGAGCTGCTTCCAATTG 4702

/ Sequence 60, Application US/09484970B
/ Patent No. 6426186
/ GENERAL INFORMATION:
/ APPLICANT: Jones, Karen A.
/ APPLICANT: Volkmuth, Wayne
/ APPLICANT: Walker, Michael G.
/ TITLE OF INVENTION: BONE REMODELING GENES
/ FILE REFERENCE: PB-0014 US
/ CURRENT APPLICATION NUMBER: US/09/484, 970B
/ CURRENT FILING DATE: 2000-01-18
/ NUMBER OF SEQ ID NOS: 172
/ SOFTWARE: PERL Program
/ SEQ ID NO 60
/ LENGTH: 13977
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. 6426186 229357.11CB1
/ LOCATION: 11721-11761, 12294, 13969
/ OTHER INFORMATION: a, t, c, g, or other
/ US-09-484-970B-60

Query Match 99.2%; Score 1290; DB 4; Length 13977;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 CGACTTTCCAGCAGTTCAGAGAGAGAGATGATGATAGGGCTTCAAGAGGGAATTGAA 60
Db 8716 CGACTTTCCAGCAGTTCAGAGAGAGAGATGATGATAGGGCTTCAAGAGGGAATTGAA 8775
QY 61 AACTAAGAAGCTGTAATCATGATGATCTTGTAGACTGTACGAATATTTCTGACAGACA 120
Db 8776 AACTAAGAAGCTGTAATCATGATGATCTTGTAGACTGTACGAATATTTCTGACAGACA 8835
QY 121 GCCTTTGAAGAGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCTCTGAGAGAG 180
Db 8836 GCCTTTGAAGAGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCTCTGAGAGAG 8895
QY 181 AGCCCAAGATGTCACTGCTTCTACGAAGCAGCTGAGAGGTCAATATCTGAGTGGGA 240
Db 8896 AGCCCAAGATGTCACTGCTTCTACGAAGCAGCTGAGAGGTCAATATCTGAGTGGGA 8955
QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAGAAGT 300
Db 8956 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAGAAGT 9015
QY 301 CCAGAACTTCAAGAGGCCAGATGAGCTGAGCTTCAAGCTGCGCAAGCTGAGGTGAT 360
Db 9016 CCAGAACTTCAAGAGGCCAGATGAGCTGAGCTTCAAGCTGCGCAAGCTGAGGTGAT 9075
QY 361 CAAGGATCCTGGCAGCCCGTGGGCGATCTCTCATTTGACTTCTCTCAAGATCACTCGA 420
Db 9076 CAAGGATCCTGGCAGCCCGTGGGCGATCTCTCATTTGACTTCTCTCAAGATCACTCGA 9135
QY 421 GAAAGTCAAGGCACCTTGGAGAGAAATTGGCCTCTGAAGAAGAACGTGAGCCAGTCAA 480
Db 9136 GAAAGTCAAGGCACCTTGGAGAGAAATTGGCCTCTGAAGAAGAACGTGAGCCAGTCAA 9195
QY 481 TGACCTTGTCTGCCAGCTTACCATTGGGCAATTCAGCTCTCAACGTATACTCAAGCAC 540
Db 9196 TGACCTTGTCTGCCAGCTTACCATTGGGCAATTCAGCTCTCAACGTATACTCAAGCAC 9255
QY 541 TCTGGAAGACCTGAACACCAAGATGAAAGCTTTCGAGGTGGCCGTGAGGACCGAGTCAG 600
Db 9256 TCTGGAAGACCTGAACACCAAGATGAAAGCTTTCGAGGTGGCCGTGAGGACCGAGTCAG 9315
QY 601 GCAGCTGATGAAGGCCACAGGACCTTGTGTCAGCATCTCAGCACTTCTTTCACAGTC 660
Db 9316 GCAGCTGATGAAGGCCACAGGACCTTGTGTCAGCATCTCAGCACTTCTTTCACAGTC 9375
QY 661 TGTCCAGGTCCTGGAGAGAGCCATCTCGCAAAAGAGTCCCTACTATATCAACCA 720

DB 9376 TGTCCAGGTCCTGGAGAGAGCCATCTGCCAACAAGTCCCTACTATATCAACCA 9435
QY 721 CGAGACTCAACAACCTTGCTGGAGACCAATCCAAATGACAGAGCTCTACAGCTTTAGC 780
DB 9436 CGAGACTCAACAACCTTGCTGGAGACCAATCCAAATGACAGAGCTCTACAGCTTTAGC 9495
QY 781 TGACCTGAATATGTGATGATTTCTAGCTTATAGAGCTGCCATGAATCCGAAGACTGCA 840
DB 9496 TGACCTGAATATGTGATGATTTCTAGCTTATAGAGCTGCCATGAATCCGAAGACTGCA 9555
QY 841 GAAGCCCTTTGCTTGATCTCTTGAGCTGTGAGCTGATGATGCTTGAGCCAGCA 900
DB 9556 GAAGCCCTTTGCTTGATCTCTTGAGCTGTGAGCTGATGATGCTTGAGCCAGCA 9615
QY 901 CAACCTCAAGCAAAATGACCAAGCCCATGATATCTGCAAGATTAATTTGTTGACCA 960
DB 9616 CAACCTCAAGCAAAATGACCAAGCCCATGATATCTGCAAGATTAATTTGTTGACCA 9675
QY 961 TATTATGACCCGCTGGAGCAAGACACAATTTGTCACGCTCCCTCTGCTGGA 1020
DB 9676 TATTATGACCCGCTGGAGCAAGACACAATTTGTCACGCTCCCTCTGCTGGA 9735
QY 1021 TATGTGCTGAAGTGGCTGCTGATGTTATGATACGGAGCAAGAGGATCCGTGT 1080
DB 9736 TATGTGCTGAAGTGGCTGCTGATGTTATGATACGGAGCAAGAGGATCCGTGT 9795
QY 1081 CCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAGTACAG 1140
DB 9796 CCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAGTACAG 9855
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACCAAGCGAGCT-GGGCC 1199
DB 9856 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACCAAGCGAGCTGGGCC 9915
QY 1200 TCCTTCTGATGATTTCTATCCAATTTCAAGACAGTGGGTGAAGTTGCATCCTTTGGGG 1259
DB 9916 TCCTTCTGATGATTTCTATCCAATTTCAAGACAGTGGGTGAAGTTGCATCCTTTGGGG 9975
QY 1260 GCAGTAACATGAGCCCAAGTGTCCGAGCTGCTTCCAATTTG 1301
DB 9976 GCAGTAACATGAGCCCAAGTGTCCGAGCTGCTTCCAATTTG 10017

RESULT 3
US-08-836-022A-10/C
; Sequence 10, Application US/08836022A
; Patent No. 6001557
; GENERAL INFORMATION:
; APPLICANT: Trustees of the University of Pennsylvania
; APPLICANT: Wilson, James M.
; APPLICANT: Fisher, Krishna J.
; APPLICANT: Chen, Shu-Jen
; APPLICANT: Weltzman, Matthew
; TITLE OF INVENTION: Improved Adenovirus Virus and
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P O Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,022A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/331,381
FILING DATE: 28-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNVFN.008PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
US-08-836-022A-10

Query Match 87.3%; Score 1135.2; DB 3; Length 19307;
Best Local Similarity 92.1%; Pred. No. 0;
Matches 1197; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 2 GACTTTCAGAGCTTCAAGAGCAGATGTACATAGGGCTTCAAGAGGATTTGAA 61
DB 5977 GATTCCAGAGCTTCAAGAGCAGATGTACATAGGGCTTCAAGAGGATTTGAA 5918
QY 62 ACTAAGAGCTGTATCATGAGTACTCTTGAGACTGTACGATATTTCTGACAGAGCAG 121
DB 5917 ACTAAGAGCTGTATCATGAGTACTCTTGAGACTGTACGATATTTCTGACAGAGCAG 5858
QY 122 CCTTTGGAAGAGCTAGAGAACTTACAGAGAGCCAGAGAGCTGCTCTGAGAGAGA 181
DB 5857 CCTTTGGAAGAGCTAGAGAACTTACAGAGAGCCAGAGAGCTGCTCTGAGAGAGA 5798
QY 182 GCCCAGAGTGTCACTGGCTTCTACGAAAGCAGGCTGAGAGGCTCATCTAGTGGAA 241
DB 5797 GCTCAGAGTGTCACTGGCTTCTACGAAAGCAGGCTGAGAGGCTCATCTAGTGGAA 5738
QY 242 AAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATATGATGAGACCCCTTGAAGACTC 301
DB 5737 AAATTGAACCTGCGCTCAGCTGATTTGGCAGAGAAAATATGATGAGACCCCTTGAAGACTC 5678
QY 302 CAGGAACCTCAAGAGGCCACGAGTGAAGCTCAAGCTGAGGCTGAGGCTGAGTATC 361
DB 5677 CAGGAACCTCAAGAGGCCACGAGTGAAGCTCAAGCTGAGGCTGAGGCTGAGTATC 5618
QY 362 AAGGATCTGCGCAGACCCGCTGGGCAATCTCTCATTTGACTCTCTCCAAGATCACCCTGAG 421
DB 5617 AAGGATCTGCGCAGACCCGCTGGGCAATCTCTCATTTGACTCTCTCCAAGATCACCCTGAG 5558
QY 422 AAAGTCAAGGCACTTTCAGAGAGAAATTCGCGCTCTGAAAGAGAAAGTGAAGCCAGCTCAT 481
DB 5557 AAAGTCAAGGCACTTTCAGAGAGAAATTCGCGCTCTGAAAGAGAAAGTGAAGCCAGCTCAT 5498
QY 482 GACCTTGCTGCGCAGCTTACCACTTTGGGCATTCAGGCTTCAACCGTATTAACCTCAGCACT 541
DB 5497 GACCTTGCACTCAGCTGACCACTGGGCATTCAGGCTTCAACCGTATTAACCTCAGCACT 5438
QY 542 CTGGAAGACTGAACACCAAGATGAAGCTTTGTCAGAGTGGCGCTGAGGAGCCAGTCAAG 601
DB 5437 TTGAAGATGTGAATACCAAGATGAAGCTTTTCAAGGTTGGCTGTGAGGAGCCGTGTGAGA 5378
QY 602 CAGCTGATGAAGCCACAGGAGCTTTGTTCAGAGATTCAGCACTTTCTTCACGCTCT 661
DB 5377 CAGCTGATGAAGCCACAGGAGCTTTGTTCAGAGATTCAGCACTTTCTTCACGCTCT 5318
QY 662 GTCCAGGCTCCCTGGAGAGAGCCATCTGCGCAACAAAGTGCCCTACTATATCAACAC 721
DB 5317 GTTCAGGCTCCCTGGAGAGAGCCATCTGCGCAACAAAGTGCCCTACTATATCAACAC 5258
QY 722 GAGACTCAACAACCTTGTGAGACCATCCCAAAATGACAGAGCTTACCAAGCTTTAGCT 781
DB 5257 GAGACTCAACAACCTTGTGAGACCATCCCAAAATGACAGAGCTTACCAAGCTTTAGCT 5198

QY 782 GACCTGAATATATGTCAGATTCTCAGCTTATAGAGCTGCCATGAAGAACTCCGAGACTGCAG 841
DB 5197 GACCTGAATATATGTCAGATTCTCAGCTTATAGAGCTGCCATGAAGAACTCCGAGACTGCAG 5138
QY 842 AAGGCCCTTGGCTTGGATCTCTTGAAGCCCTGTCAGCTGCATGTGATGCCCTTGGACCAAGCAC 901
DB 5137 AAGGCCCTTGGCTTGGATCTCTTGAAGCCCTGTCAGCTGCATGTGATGCCCTTGGACCAAGCAC 5078
QY 902 AACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAATTAATTAATTTGAACCACT 961
DB 5077 AACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAATTAATTAATTTGAACCACT 5018
QY 962 ATTATGACCGCCTGAGCAAGACCAACAATTGTCACGCTCCCTCTCTGCTGGAT 1021
DB 5017 ATTATGACCGCCTGAGCAAGACCAACAATTGTCACGCTCCCTCTCTGCTGGAT 4958
QY 1022 ATGTGTCGAAGCTGCTGCTGAATGTTATGATACGGAAGCAAGGAGGATCCGCTGC 1081
DB 4957 ATGTGTCGAAGCTGCTGCTGAATGTTATGATACGGAAGCAAGGAGGATCCGCTGC 4898
QY 1082 CTGTCTTTAAACTGGCATCATTTCCCTGTGTAAGACATTTGGAAGACAGTACAGA 1141
DB 4897 CTGTCTTTAAACTGGCATCATTTCCCTGTGTAAGACATTTGGAAGACAGTACAGA 4838
QY 1142 TACCTTTCAAGCAAGTGGCAAGTCAACAGATTTTGTGACCGACGAGCTGGGCTC 1201
DB 4837 TACCTTTCAAGCAAGTGGCAAGTCAACAGATTTTGTGACCGACGAGCTGGGCTC 4778
QY 1202 CTCTGTCATGATTTCTATCCAAATTCACAGACAGTGGTGAAGTTGCATCCCTTGGGGGC 1261
DB 4777 CTCTGTCATGATTTCTATCCAAATTCACAGACAGTGGTGAAGTTGCATCCCTTGGGGGC 4718
QY 1262 AGTAACATGAGCCCAAGTGTCCGAGCTGCTTCCAATTTG 1301
DB 4717 AGTAACATGAGCCCAAGTGTCCGAGCTGCTTCCAATTTG 4678

RESULT 4

US-09-427-048A-10/c
Sequence 10, Application US/09427048A

Patent No. 6203975

GENERAL INFORMATION:

APPLICANT: Trustees of the University of Pennsylvania

Wilson, James M.

Fisher, Krishna J.

Chen, Shu-Jen

Weitzman, Matthew

TITLE OF INVENTION: Improved Adenovirus Virus and

Methods of Use Thereof

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson

STREET: Spring House Corporate Cntr, P O Box 457

CITY: Spring House

STATE: Pennsylvania

COUNTRY: USA

ZIP: 19477

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/427,048A

FILING DATE: 21-Oct-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/836,022

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Bak, Mary E.

REGISTRATION NUMBER: 31,215

REFERENCE/DOCKET NUMBER: GNPVN.008PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-427-048A-10

Query Match 87.3%; Score 1135.2; DB 3; Length 19307;
Best Local Similarity 92.1%; Pred. No. 0;
Matches 1197; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 2 GACTTCCAGCAGTTCAGAGAGCAAGCATGTATAGAGCCCTTCAAGAGGAATTGAAA 61
DB 5977 GATTTCAGCAGAGTTCAGAGAGCAAGATGATATACATAGGGCTTCAAGAGGAATTGAAA 5918
QY 62 ACTAAGAACTGTATCATAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAG 121
DB 5917 ACTAAGAACTGTATCATAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAG 5858
QY 122 CCTTGGAAAGACTAGAGAACTCTACAGAGAGCCCAAGAGAGCTGCTCCTGAGAGAGA 181
DB 5857 CCTTGGAAAGACTAGAGAACTCTACAGAGAGCCCAAGAGAGCTGCTCCTGAGAGAGA 5798
QY 182 GCCCAGATGTCACTCGGCTTCTACGAAGACAGGCTGAGAGAGTCAATACTGAGTGGAA 241
DB 5797 GCTCAGATGTCACTCGGCTTCTACGAAGACAGGCTGAGAGAGTCAATACTGAGTGGAA 5738
QY 242 AAATTGAACCTGCACTCCGCTGACTGCGCAGAGAGAAATAGATGAGACCTTGAAAGACTC 301
DB 5737 AAATTGAACCTGCACTCCGCTGACTGCGCAGAGAGAAATAGATGAGACCTTGAAAGACTC 5678
QY 302 CAGAACTTCAAGAGGCCAGGATGAGCTGAGACCTCAAGCTGCGCAAGCTGAGGTGATC 361
DB 5677 CAGAACTTCAAGAGGCCAGGATGAGCTGAGACCTCAAGCTGCGCAAGCTGAGGTGATC 5618
QY 362 AAGGATCTTGGCAGAGCCGCTGAGCTCTCTCATTTGACTCTCTCCAAAGATCACTCGAG 421
DB 5617 AAGGATCTTGGCAGAGCCGCTGAGCTCTCTCATTTGACTCTCTCCAAAGATCACTCGAG 5558
QY 422 AAAGTCAAGGCACTTCGAGAGAGAAATGCGCTCTGAAGAGAGAGCTGAGCCAGCTCAAT 481
DB 5557 AAAGTCAAGGCACTTCGAGAGAGAAATGCGCTCTGAAGAGAGAGCTGAGCCAGCTCAAT 5498
QY 482 GACCTTGTCCGCCAGCTTACCACTTTGGGATTCAGCTCTCAACCGTATAACCTCAGCACT 541
DB 5497 GACCTTGTCCGCCAGCTTACCACTTTGGGATTCAGCTCTCAACCGTATAACCTCAGCACT 5438
QY 542 CTGGAAGACCTGAACACCAAGATGAGAACTTCTGCAAGTGGCCCTCGAGAGACCGAGTCAGG 601
DB 5437 TTGGAAGATCTGAATAACAGATGAGGCTTCTACAGGTGGCTGTGAGAGACCGTGTGAGA 5378
QY 602 CAGCTGATGAAGCCCAAGGACTTTGGTCCAGCATCTCAGCACTTCTTTCCAGCTCT 661
DB 5377 CAGCTGATGAAGCCCAAGGACTTTGGTCCAGCATCTCAGCACTTCTTTCCAGCTCTCA 5318
QY 662 GTCCAGGCTCCCTGGAGAGAGGCACTCTGCCAAACAAAGTGCCTACTATATCAACCAAC 721
DB 5317 GTCCAGGCTCCCTGGAGAGAGGCACTCTGCCAAACAAAGTGCCTACTATATCAACCAAC 5258
QY 722 GAGACTCAACCAACTGTCTGGGACATCCCAAAATGACAGAGCTTACCAAGCTTTAGCT 781
DB 5257 GAGACTCAACCAACTGTCTGGGACATCCCAAAATGACAGAGCTTACCAAGCTTTAGCT 5198
QY 782 GACCTGAATATATGTCAGATTCTCAGCTTATAGAGCTGCCATGAAGAACTCCGAGAGCTGCAG 841
DB 5197 GACCTGAATATATGTCAGATTCTCAGCTTATAGAGCTGCCATGAAGAACTCCGAGAGCTGCAG 5138

| | | | |
|----|------|--|------|
| QY | 842 | AAGGCCCTTGGCTTGGATCTCTTGAGCCCTGTACGTCATGTGATGCCCTTGACCAAGCAC | 901 |
| Db | 5137 | AAGGCCCTTGGCTTGGATCTCTTGAGCCCTGTACGTCATGTGATGCCCTTGACCAAGCAC | 5078 |
| QY | 902 | AACCTCAAGCAAAATGACCAAGCCCATGATATCCTGAGATTATTAATTGTTGACCACT | 961 |
| Db | 5077 | AACCTCAAGCAAAATGACCAAGCCCATGATATCCTGAGATTATTAATTGTTGACTACA | 5018 |
| QY | 962 | ATTATGACCGCCTGGAGCAAGACACAACAATTGGTCAACGTCCCTCTCGCTGGAT | 1021 |
| Db | 5017 | ATTATGATCGTCTGGAGCAAGACACAACAATCTGTCAATGTCCCTCTCGTGGAT | 4958 |
| QY | 1022 | ATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGACGAACAGGAGATCCGTGC | 1081 |
| Db | 4957 | ATGTGCTCAACCTGGCTTCTCAATGTTTATGATACGGACGAACAGGAGATCCGTGC | 4898 |
| QY | 1082 | CTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAGTACAGA | 1141 |
| Db | 4897 | CTGTCTTTTAAACTGGCATCATTTCTCTGTGTAAGCACACTTGAAGACAAGTACAGA | 4838 |
| QY | 1142 | TACCTTTTCAAGCAAGTGGCAAGTTCACACAGATTTTGTGACCAGCGCAGCGTGGCCTC | 1201 |
| Db | 4837 | TACCTTTTCAAGCAAGTGGCAAGTTCACACTGGCTTTTGTGACCAGCGTAGCGTCTT | 4778 |
| QY | 1202 | CTTCTGCATGATTCATCCCAAATTCCAAGACAGTGGGTGAAGTTGCATCCTTTGGGGGC | 1261 |
| Db | 4777 | CTTCTGCATGATTCATCCCAAATCCCAAGACAGTGGGTGAAGTTGCTTCTTTGGGGGC | 4718 |
| QY | 1262 | AGTAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAAATTTG 1301 | |
| Db | 4717 | AGTAACATTGAGCCCAAGTGTGAGAGAGCTGCTTCCAAATTTG 4678 | |

RESULT 5

```

? Sequence 7, Application US/09091501B
? Patent No. 6518413
? GENERAL INFORMATION:
? APPLICANT: Tinsley, Jonathon M
? APPLICANT: Davies, Kay E
? TITLE OF INVENTION: Uterophin gene expression
? FILE REFERENCE: 620-42
? CURRENT APPLICATION NUMBER: US/09/091,501B
? CURRENT FILING DATE: 1998-06-18
? PRIOR APPLICATION NUMBER: PCT/GB96/03156
? PRIOR FILING DATE: 1996-12-19
? PRIOR APPLICATION NUMBER: GB 9525962.8
? PRIOR FILING DATE: 1995-12-19
? PRIOR APPLICATION NUMBER: GB 9615797.9
? PRIOR FILING DATE: 1996-07-26
? PRIOR APPLICATION NUMBER: GB 9622174.2
? PRIOR FILING DATE: 1996-10-24
? NUMBER OF SEQ ID NOS: 15
? SOFTWARE: Patentin Ver. 2.1
? SEQ ID NO 7
? LENGTH: 6045
? TYPE: DNA
? ORGANISM: Artificial Sequence
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (11)..(6037)
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence: Chimeric
? FEATURE:
? NAME/KEY: misc feature
? LOCATION: (724)..(758)
? OTHER INFORMATION: Precise residue is left open
? US-09-091-501B-7

```

| | | | | |
|-----------------------|-----------------|---------------------|-----------|--------------|
| Query Match | 42.6% | Score 553.8; | DB 4; | Length 6045; |
| Best Local Similarity | 64.7%; | Pred. No. 3.8e-169; | | |
| Matches 844; | Conservative 0; | Mismatches 452; | Indels 9; | Gaps 1; |

| | | | |
|----|------|---|------|
| QY | 2 | GACTTTCACGACGTCAGAGACAGAACGATGTACATAGAGCCCTTCAAGAGGGAAATTGAAA | 61 |
| Db | 3506 | GATGTTCCAGCCTTACAGCTCCAGTATGACCATTGTGAAGGCCCTGAGACGGGAGTTAAAG | 3565 |
| QY | 62 | ACTAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGCACAGACGAG | 121 |
| Db | 3566 | GAGAAAGATATTCTGTCTCTGAATGCTGTGCACCAGGCCCGAGTTTCTTGCTGATCAG | 3625 |
| QY | 122 | CCTTTGGAAGGACTAGAGAACT-----CTACAGGAGCCACAGAGAGCTGCCTCT | 172 |
| Db | 3626 | CCAATTGAGGCCCTGAGAGCCAGAAAGAACTCAATCAAAAACAGAAATTAATCTCT | 3685 |
| QY | 173 | GAGGAGAGAGCCCGAAGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT | 232 |
| Db | 3686 | GAGGAGAGAGCCCAAAAGATTGCCAAAGCCATGCGCAACAGTCTTCTGAGTCAAAAGAA | 3745 |
| QY | 233 | GAGTGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCT | 292 |
| Db | 3746 | AAATGGAAAACTTAATGCTGTAACTAGCAATTGGCAAAAAGCAAGTGACAGGCAATTG | 3805 |
| QY | 293 | GAAGACTCCAGGAACCTTCAAGAGCCACGATGAGCTGGAACCTCAAGCTGCCCAAGCT | 352 |
| Db | 3806 | GAGAACTCAGAGACCTGCAGGAGCTATGATGACCTGGACGCTGACATGAGAGAGCA | 3865 |
| QY | 353 | GAGGTATCAAGGATCCTGGCAGCCCGTGGCGCATCTCTCATTTGACTCTCTCCAAGAT | 412 |
| Db | 3866 | GAGTCGTGCGGAATGCTGGAAGCCCGTGGAGACTTACTCATTTGACTCGCTGCAGAT | 3925 |
| QY | 413 | CACCTGAGAAAGTCAAGGCACTTGAGAGAAATTGCGCCTCTGAAGAAGACGTGAGC | 472 |
| Db | 3926 | CACATTGAAAAAATCATGGCAATTAGAGAGAAATTGCAACCAATCACTTTAAAGTTAAA | 3985 |
| QY | 473 | CAGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAATTGAGCTCTGACGTATTAAC | 532 |
| Db | 3986 | ACGGTGAATGATTATTCAGTCAAGCTGTCTCCACTGACCTGCATCCCTCTCTAAAGATG | 4045 |
| QY | 533 | CTCAGCACTTGGAAGACTGAAACACAGATGGAAGCTTCTGAGGTGCCGTGAGGAC | 592 |
| Db | 4046 | TCTCGCCAGCTAGATGACCTTAATATGCGATGGAACCTTTACAGGTTTCTGTGATGAT | 4105 |
| QY | 593 | CGAGTCAGGACGCTGCATGAAGCCCAAGGACCTTTGCTCCAGCATCTCAGCACTTCTT | 652 |
| Db | 4106 | CGCCTTAAACAGCTTCAGGAAGCCCAAGAGATTTGGAACCATCTCTCAGACATTTCTC | 4165 |
| QY | 653 | TCCAGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTGCCAACAAGTGCCTACTAT | 712 |
| Db | 4166 | TCTACGTCACTCCAGCTGCCGTGGCAAGATCCATTTTCACATATAAAGTCCCTATTAC | 4225 |
| QY | 713 | ATCAACCAAGAGACTCAAAACAATTGCTGGGACCATCCCAAAATGACAGAGCTTACCAG | 772 |
| Db | 4226 | ATCAACCATCAAAACACAGACCACTGTTGGGACCATCTTAAATGACCGAACTCTTTCAA | 4285 |
| QY | 773 | TCTTAGCTGACCTGAATATGTCAAGTCTCAGCTTATAGGACTGCCATGAAACTCCGA | 832 |
| Db | 4286 | TCCCTTGCTGACCTGAATATGTACGTTTCTTCTGCTTACCCTACCGTACAGCAATCAAAATCCGA | 4345 |
| QY | 833 | AGACTGAGAAAGGCCCTTGTGATCTCTTGAGCCTGTACGCTGCATGTGATGCCCTG | 892 |
| Db | 4346 | AGACTCAAAAAAGCACTATGTTGGATCTCTTAGAGTTGAGTACAAACAAATGAAATTTTC | 4405 |
| QY | 893 | GACCAAGCAAACTCTAAGCAAAATGACCAAGCCCATGGATATCTGACAGATTAATTAATTGT | 952 |
| Db | 4406 | AAACAGACAAGTTGAACCAAAATGACCAAGCTCCTCAGTGTTCCAGATGTATCAACTGT | 4465 |
| QY | 953 | TTGACCACTATTTATGACCGCTGGAGCAAGACAACAATTGTTGTTCAAGCTCCCTCTC | 1012 |
| Db | 4466 | CTGACAAACAATTATGATGAGACTTGAGCAAAATGCAATAAGGACCTGTTCAACGTTTCCACTC | 4525 |
| QY | 1013 | TGCGTGATATGTGTCTGAAGTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGG | 1072 |
| Db | 4526 | TGTGTTGATATGTGTCTCAATTGTTGTTGCTCAATGCTATGACACGGGTGCAACTGGAAAA | 4585 |

Db 4526 TGTGTGATATGTGTCTCAATTGGTTGCTCAATGCTTAATGACACGGGTCCAACTGGAAA 4585

QY 1073 ATCCGTGCTCTGCTTTTAAACTGGCATTCATTCCTGTGTAAAGCACATTTGGAAGAC 1132
| | | | |
Db 4586 ATTAGAGTGACAGAGCTGAGATTTGATTAATGTCTCTCCAAAGGCTCTTGGAAGAA 4645
QY 1133 AAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAAGGATTTTGTGACCAAGCGCAGG 1192
| | | | |
Db 4646 AATAACAGATATCTCTTTAAGAAAGTGGCGGGCCGACAGAAATGTGTGACCAAGCGCAG 4705
QY 1193 CTGGGCTCTCTTCGATGATTTCTATCCAAATTCACAGACAGTGGTGAAGTTCATCC 1252
| | | | |
Db 4706 CTGGGCTCTTACTTCAATGATCCATCCAGATCCCGGACGTAAGTGAAGTACAGCT 4765
QY 1253 TTTGGGGGAGTAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAA 1297
| | | | |
Db 4766 TTTGGAGGAGTAATATTGAGCTTAGTGTCCAGCTGCTTCCAA 4810

RESULT 6

US-09-091-501B-9
; Sequence 9, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 10320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(10312)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Full length
; OTHER INFORMATION: utrophin construct
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (724)..(758)
; OTHER INFORMATION: Precise residue is left open
US-09-091-501B-9

Query Match 42.6%; Score 553.8; DB 4; Length 10320;
Best Local Similarity 64.7%; Pred. No. 5.3e-169;
Matches 844; Conservative 0; Mismatches 452; Indels 9; Gaps 1;

QY 2 GACTTTCACAGCAGTTCAGAGCAAGATGTACATAGGGCTTCAAGAGGGAATTGAAA 61
| | | | |
Db 7781 GATGTTCCAGCCTTACAGCTCCAGTATGACCATTTAGAGCCCTGAGACGGAGTTAAAG 7840
QY 62 ACTAAGAACTGTATCATGAGTACTCTTGAGACTGTAGCAATATTTCTGACAGAGCAG 121
| | | | |
Db 7841 GAGAAAGAAATATCTGTCTGATGCTGTGACAGAGCCGAGTTTCTTGCTGATCAG 7900
QY 122 CCTTTGGAAGACTAGAGAACT-----CTACAGAGAGCCAGAGAGCTGCTCCT 172
| | | | |
Db 7901 CCAATTGAGGCCCTGAGAGAGCCAGAGAACTACATCAAAAAACAGAAATTAATCCT 7960
QY 173 GAGGAGAGAGCCAGAGATGTCACTTCGCTTCTACGAAAGAGAGCTGAGAGGTCAATACT 232
| | | | |
Db 7961 GAGGAGAGAGCCAGAGATGTCCCAAGCCATGCGCAAGAGTCTTCTGAAAGTCAAGAA 8020

QY 233 GAGTGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATGATGAGACCCCTT 292
| | | | |
Db 8021 AATGGGAAAGCTTAATGCTGTACTAGCAATTGGCAAAAGCAAGTGAACAAGCATTTG 8080
QY 293 GAAAGACTCCAGAACTTCAAGAGCCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 352
| | | | |
Db 8081 GAGAACTCAGAGAGCTGAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 8140
QY 353 GAGGTGATCAAGGATCTGGCAGCCCGTGGGCGATCTCTCATTTGATCTCTCCAGAT 412
| | | | |
Db 8141 GAGTCCGTCGGGAATGGCTGGAAGCCCGTGGGAGACTTACTCATTTGACTGCTGCAGAT 8200
QY 413 CACCTCGAGAAAGTCAAGGCACTTCGAGAGAAATGCGCCTGTAAGAGAGAGTGAAGC 472
| | | | |
Db 8201 CACATTTGAAAAATCATGCGCATTTAGAGAGAAATGCAACCAATCAACTTAAGTTAAA 8260
QY 473 CACGTCAATGACCTTGTCTGCGCAGCTTACCACTTTGGGCATTCAGCTTCACCCGTATAAC 532
| | | | |
Db 8261 ACGGTGAATGATTTATTCAGTCAAGCTGTCTCCACTTGAACCTGCATCCCTCTTAAGATG 8320
QY 533 CTGACCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCCGTCGAGAC 592
| | | | |
Db 8321 TCTGCGCAGCTAGATGACCTTAATATGCGATGGAAGCTTTACAGGTTTCTGTGATGAT 8380
QY 593 CGAGTCAGGAGCTGATGAAGCCCAAGGAGCTTTGTCCAGCATCTCAGCACTTTCTT 652
| | | | |
Db 8381 CGCCTTAAACAGCTTCAGGAGAGCCCAAGAGATTTTGAACCATCTCTCAGCACTTTCTC 8440
QY 653 TCACGCTGTCTCAGGCTGCTGCGGAGAGAGAGCAATCTGCCAAACAAGTGCCTACTAT 712
| | | | |
Db 8441 TCTAGCTGAGTCCAGCTGCGCTGCGCAAGATTCATTAATATTAAGTGCCTACTAT 8500
QY 713 ATCAACCAAGAGACTCAACCACTTGTCTGGGACATCCCAAAATGACAGAGCTTACCAG 772
| | | | |
Db 8501 ATCAACCAATCAACCAAGAGCAACCACTGTTGGGACATCTTAATATGAACCAAGCTTTCAA 8560
QY 773 TCTTTAGCTGACCTGAATTAATGTCAGATTTTCAGCTTATAGGAGTGCATGAATCCGA 832
| | | | |
Db 8561 TCCCTGCTGACCTGAATTAATGTCAGTCTTTTCTGCTCAACCTGACAGCAATCAAAATCCGA 8620
QY 833 AGACTGACAGAGGCCCTTGTCTGATCTCTTGAAGCTGTGAGCTGATGATGCTTGTG 892
| | | | |
Db 8621 AGACTACAAAAAGACATGTTGATCTCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 8680
QY 893 GACCAACCAACCTCAAGCAAAATGACCAAGCCATGATATCTGACAGATTAATTAATGT 952
| | | | |
Db 8681 AAACAGCAAGTGAACCAAAATGACCAAGCTCTCAAGTGTCCAGATGTCATCAACTGT 8740
QY 953 TTGACCACTATTATGACCGCCTGAGCAAGAGCAACAATTTGTCAAGCTCCCTCTC 1012
| | | | |
Db 8741 CTGACAACACTTATGATGAGCTTGAAGCAATGCAAGAGCTGTCAAGCTTCCACTC 8800
QY 1013 TGGGTGATATGTGTCTGAAGTGTGCTGCTGATGTTTATGATACGGAGCAAGAGAGG 1072
| | | | |
Db 8801 TGTGTTGATATGTGTCTCAATGCTGCTCAATGCTATGACAGGCTGCAACTGGAAGAA 8860
QY 1073 ATCCGTGCTCTGTCTTTTAAACTGCGATCATTTCCCTGTGTAAGACACATTTGGAAGAC 1132
| | | | |
Db 8861 ATTAGAGTCAAGAGTCTGAAGATTTGATTAATGTCTCTCCAAAGCTCTCTTGAAGAA 8920
QY 1133 AAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAAGAGATTTTGTGACCAAGCGCAGG 1192
| | | | |
Db 8921 AATACAGATATCTCTTTAAGAAAGTGGCGGGCCGACAGAAATGTGTGACCAAGCGCAG 8980
QY 1193 CTGGGCTCTCTTCTGATGATTTCTATCCAAATTTCCAAAGAGAGTGGGTGAAGTTCATCC 1252
| | | | |
Db 8981 CTGGGCTGTTACTTCAATGATGCCATCCAGATCCCGGAGAGCTAGGTGAAGTACAGCT 9040
QY 1253 TTTGGGAGAGTAACATTGAGCCCAAGTGTCCGAGAGCTTCCAA 1297
| | | | |
Db 9041 TTTGAGAGAGTAATATTGAGCTTAGTGTTCGAGAGCTTCCAA 9085

RESULT 7

US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PTZgpt-Fls
; US-08-232-463-14

Query Match 3.8%; Score 50; DB 1; Length 7218;

Best Local Similarity 4.2%; Pred. No. 2.1e-05;

Matches 14; Conservative 191; Mismatches 131; Indels 0; Gaps 0;

QY 18 AGAAGCAGACGATGTACATAGGGCTTCAAGAGGGAATTGAAACTAAAGAACTGTAA 77
DB 1381 RRR 1322
QY 78 TCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCTTTGGAAGACTAG 137
DB 1321 RRR 1262
QY 138 AGAAACTCTACAGAGCCAGAGAGCTGCCTCCTGAGAGAGAGCCAGAAATGTCACTC 197
DB 1261 RRR 1202
QY 198 GGCTTCTACGAAAGCAGGCTGAGAGAGTCAATACTGAGTGGAAAAATGAACCTGCACT 257
DB 1201 RRR 1142
QY 258 CCGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAGACTCCAGAACTTCAAGAGG 317
DB 1141 RRR 1082

QY 318 CCACGATGAGCTGAGACTCAAGCTGCCGCAAGCTG 353

DB 1081 RRRRRRRRRRRRRRRATCGCAAGCTTCCTCGACCTG 1046

RESULT 8

US-09-668-313A-10
; Sequence 10, Application US/09668313A
; Patent No. 6503756
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Susan M. Freier
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF SYNTAXIN 4 INTERACTING PROTEIN EXPRES.
; FILE REFERENCE: RTS-0127
; CURRENT APPLICATION NUMBER: US/09/668,313A
; CURRENT FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 247
; SEQ ID NO 10
; LENGTH: 2574
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (218)...(1891)
; US-09-668-313A-10

Query Match 3.4%; Score 44.2; DB 4; Length 2574;

Best Local Similarity 54.7%; Pred. No. 0.00086;

Matches 88; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 674 TGGAGAGAGCCATCTGCCAAACAAGTGCCCTACTATATCAACAGAGAGCTGAACA 733
DB 1733 TGGAGAGAGCTTACACAGCAGATGAATCAAGTACTTCATCAACACGTCAGACAGACC 1792
QY 734 ACTTGCTGGAGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTAGCTGACCTGAATAAT 793
DB 1793 ACGTCTGATCCACCCCGTGTGAGCGCCCTGAACCTGTCTGTGACAGAGAGTGAA 1852
QY 794 GTCAAGTTCTCAGCTTATAGAGCTGCCATGAACCTCCGAAG 834
DB 1853 GAGAGCTGCCAGAGAGCTTAACAGACCCGAAAGCTGATG 1893

RESULT 9

US-09-620-312D-69

; Sequence 69, Application US/09620312D

; Patent No. 6569662

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Liu, Chenghua

; APPLICANT: Asundi, Vinod

; APPLICANT: Zhang, Jie

; APPLICANT: Ren, Feiyan

; APPLICANT: Chen, Rui-hong

; APPLICANT: Zhao, Qing A.

; APPLICANT: Wehrman, Tom

; APPLICANT: Xue, Aidong J.

; APPLICANT: Yang, Yonghong

; APPLICANT: Wang, Jian-Rui

; APPLICANT: Zhou, Ping

; APPLICANT: Ma, Yungqing

; APPLICANT: Wang, Dunrui

; APPLICANT: Wang, Zhiwei

; APPLICANT: John Tillinghast

; APPLICANT: Drmanac, Radoje T.

; TITLE OF INVENTION: No. 6569662el Nucleic Acids and

; FILE REFERENCE: 784CIP2B

; CURRENT APPLICATION NUMBER: US/09/620,312D

; CURRENT FILING DATE: 2000-07-19

; PRIOR APPLICATION NUMBER: 09/552,317

;; PRIOR FILING DATE: 2000-04-25
;; PRIOR APPLICATION NUMBER: 09/488,725
;; PRIOR FILING DATE: 2000-01-21
;; NUMBER OF SEQ ID NOS: 1105
;; SOFTWARE: pt_fl_genes Version 1.0
;; SEQ ID NO 69
;; LENGTH: 1690
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (128)..(1522)
US-09-620-312D-69

Query Match 3.3%; Score 42.8; DB 4; Length 1690;
Best Local Similarity 47.9%; Pred. No. 0.0019;
Matches 156; Conservative 0; Mismatches 167; Indels 3; Gaps 1;

QY 301 CCAGAACTTCAAGAGCCAGATGAGCTGACCTCAAGCTGCCCAAGCTGAGTGAT 360
DB 55 CAAGAGATTGCACCAAGTGGCGCAGACCTGACGACGAGCTGGCATGGTTGAGAGCG 114
QY 361 CAAGGA---TCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTTCCAGATCACCT 417
DB 115 GCTGCCACTGGCCATGACAGACAGGCAACGGTTGACAGCGCGTCCAGACGACAT 174
QY 418 CGAGAAAGTCAAGGCACTTCGAGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCAGT 477
DB 175 CAAAAGAACCAAGGCGCTGCGCGGAGATCCAGGCGCATGGCGCGCTGAGAGAGGT 234
QY 478 CAATGACCTTGCTCGCCAGCTTACCACTTTGGGCACTTCAACCGTATTAACCTCAG 537
DB 235 GCTGAGCGCGCGCGCGCTGCTGCTGCGCAGCCCGGAGGCAAGGCAAGTGGCGCG 294
QY 538 CACTCTGAAGACCTGAACACCAAGATGGAAGCTTCTGACAGTGGCCGTCGAGACCGAGT 597
DB 295 GGGCCTGAGCAGCTGCAGAGCGCTGGCGCGGACTGCGGAGGCGCTGCCGAGCAAGGCA 354
QY 598 CAGGAGCTGCATGAAGCCCAAGG 623
DB 355 GCAGGTGCTGAGCGCGCTTCCAGG 380

RESULT 10

US-09-368-590-1
; Sequence 1, Application US/09368590
; Patent No. 6187563
; GENERAL INFORMATION:
; APPLICANT: Solimena, Michele
; TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
; TITLE OF INVENTION: AUTOANTIGENS OF AUTOIMMUNE DISEASES
; FILE REFERENCE: 101918-200 (OCR-941)
; CURRENT APPLICATION NUMBER: US/09/368,590
; EARLIER FILING DATE: 1999-08-04
; EARLIER APPLICATION NUMBER: 60/095,657
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 7812
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(6879)
; NAME/KEY: unsure
; LOCATION: (100)..(102)
; NAME/KEY: unsure
; LOCATION: (1021)..(1023)
; NAME/KEY: unsure
; LOCATION: (2266)..(2268)
US-09-368-590-1

Query Match 3.3%; Score 42.8; DB 3; Length 7812;
Best Local Similarity 47.9%; Pred. No. 0.0049;
Matches 156; Conservative 0; Mismatches 167; Indels 3; Gaps 1;

QY 301 CCAGAACTTCAAGAGCCAGATGAGCTGACCTCAAGCTGCCCAAGCTGAGTGAT 360
DB 3639 CAAGAGTTGCACCAAGTGGCGCAGACCTGACGACGAGCTGGCATGGTTCAAGAGCG 3698
QY 361 CAAGGA---TCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTTCCAGATCACCT 417
DB 3699 GCTGCCACTGGCCATGACAGACAGCGGCAACGGTTTGCAGGCGGTTCCAGACGACAT 3758
QY 418 CGAGAAAGTCAAGGCACTTCAGAGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCAGT 477
DB 3759 CAAAAGAACCAAGGCGCTGCGCGGAGATCCAGGCGCATGGCGCGCTGAGAGAGT 3818
QY 478 CAATGACCTTGCTCGCCAGCTTACCACTTTGGGCACTTCACTCTCACCGTATTAACCTCAG 537
DB 3819 GCTGAGCGCGCGCGCTGCGCTGCTGCGCAGCCCGGAGGCAAGGCAAGTGGCGCG 3878
QY 538 CACTCTGAAGACCTGAACACCAAGATGGAAGCTTCTGACAGTGGCCGTCGAGACCGAGT 597
DB 3879 GGGCCTGAGCAGCTGCAGAGGCGCTGGCGCGGACTGCGGAGGAGGCTGCCGAGCAAGGCA 3938
QY 598 CAGGAGCTGCATGAAGCCCAAGG 623
DB 3939 GCAGGTGCTGAGCGCGCTTCCAGG 3964

RESULT 11

US-08-425-069-3
; Sequence 3, Application US/08425069
; Patent No. 5728810
; GENERAL INFORMATION:
; APPLICANT: Lewis, Randolph V.
; APPLICANT: Xu, Ming
; APPLICANT: Himan, Michael B.
; TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
; TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
; TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 No. 5728810th Washington Street
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/425,069
; FILING DATE: 19-APR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1447-106P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1995 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO

IMMEDIATE SOURCE:
CLONE: p6B
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1785
US-08-425-069-3

Query Match 3.0%; Score 38.6; DB 1; Length 1995;
Best Local Similarity 44.0%; Pred. No. 0.049;
Matches 164; Conservative 0; Mismatches 209; Indels 0; Gaps 0;

QY 11 GCAGTTCAGAGCAGAACGATGTACATAGGCGCTTCAGAGGGAATTGAAACTAAAGAA 70
DB 731 GCAGTGCAGCTGCAGACGCCGACAGACCTGACACACAGAGCCCGAGAGATATGAC 790
QY 71 CCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGACAGCCTTTGAA 130
DB 791 CAGGACACAAAGAACCATCTGACCCGGTAGTGCCGCTGCAGACAGACGCCGCGACAG 850
QY 131 GGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCCTCCTGAGAGAGAGCCAGAAT 190
DB 851 GACCTGAGAGATATGGCCCTGGACACAAAGAGACCCGAGAGATATGACCAAGCAAG 910
QY 191 GTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCACTACTGAGTGGGAAATTTGAAC 250
DB 911 GACCATCTGAGACAGCAGAGTGCAGACAGACAGCCGACAGAGACCTTGACACAAAGAT 970
QY 251 CTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTGAAGACTCCAGAACTT 310
DB 971 TAGGAGTTATGAGACCAAGACACAAAGTCCAGAGAGATATGACCAAGACACAAAGTTC 1030
QY 311 CAAGAGCCACCGATGAGCTGAGCTCAAGCTGCCGCCAGCTGAGGTATCAAGGATCC 370
DB 1031 CAGGAGATATGAGACCAAGTGTAGTCTGACAGACAGCCGACAGAGAGACCTTGACAC 1090
QY 371 TGGCAGCCCGTGG 383
DB 1091 AAGGACCAAGAGG 1103

RESULT 12

US-08-317-844B-3

Sequence 3, Application US/08317844B

Patent No. 5989894

GENERAL INFORMATION:

APPLICANT: Lewis, Randolph V.

APPLICANT: Xu, Ming

APPLICANT: Hlman, Michael B.

TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK

TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL

TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF

NUMBER OF SEQUENCES: 62

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Birch, Stewart, Kolasch & Birch

STREET: 301 No. 5989894th Washington Street

CITY: Falls Church

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22046

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/317,844B

FILING DATE: 04-OCT-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Murphy Jr., Gerald M

REGISTRATION NUMBER: 28,977

REFERENCE/DOCKET NUMBER: 1447-105P

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 241-1300

TELEFAX: (703) 241-2848

TELEX: 248345

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1995 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

IMMEDIATE SOURCE:

CLONE: p6B

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1785

US-08-317-844B-3

Query Match 3.0%; Score 38.6; DB 2; Length 1995;
Best Local Similarity 44.0%; Pred. No. 0.049;
Matches 164; Conservative 0; Mismatches 209; Indels 0; Gaps 0;

QY 11 GCAGTTCAGAGCAGAACGATGTACATAGGCGCTTCAGAGGGAATTGAAACTAAAGAA 70
DB 731 GCAGTGCAGCTGCAGACGCCGACAGACCTGACACACAGAGCCCGAGAGATATGAC 790
QY 71 CCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGACAGCCTTTGAA 130
DB 791 CAGGACACAAAGAACCATCTGACCCGGTAGTGCCGCTGCAGACAGACGCCGCGACAG 850
QY 131 GGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCCTCCTGAGAGAGAGCCAGAAT 190
DB 851 GACCTGAGAGATATGGCCCTGGACACAAAGAGACCCGAGAGATATGACCAAGCAAG 910
QY 191 GTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCACTACTGAGTGGGAAATTTGAAC 250
DB 911 GACCATCTGAGACAGCAGAGTGCAGACAGACAGCCGACAGAGACCTTGACACAAAGAT 970
QY 251 CTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTGAAGACTCCAGAACTT 310
DB 971 TAGGAGTTATGAGACCAAGACACAAAGTCCAGAGAGATATGACCAAGACACAAAGTTC 1030
QY 311 CAAGAGCCACCGATGAGCTGAGCTCAAGCTGCCGCCAGCTGAGGTATCAAGGATCC 370
DB 1031 CAGGAGATATGAGACCAAGTGTAGTCTGACAGACAGCCGACAGAGAGACCTTGACAC 1090
QY 371 TGGCAGCCCGTGG 383
DB 1091 AAGGACCAAGAGG 1103

RESULT 13

US-09-220-132-24

Sequence 24, Application US/09220132

Patent No. 6506607

GENERAL INFORMATION:

APPLICANT: Shayjan, Andrew W.

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT

TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CP

FILE REFERENCE: 07334-074001

CURRENT APPLICATION NUMBER: US/09/220,132

PRIOR FILING DATE: 1998-12-23

PRIOR APPLICATION NUMBER: US 60/079,303

PRIOR FILING DATE: 1998-03-25

PRIOR APPLICATION NUMBER: US 60/068,821

PRIOR FILING DATE: 1997-12-24

NUMBER OF SEQ ID NOS: 191

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 24

LENGTH: 7672

TYPE: DNA

ORGANISM: Homo sapiens

US-09-220-132-24

| | | | | |
|---------------------------|--------|-----------------|-----------|--------------|
| Query Match | 3.08; | Score 38.4; | DB 4; | Length 7672; |
| Best Local Similarity | 47.28; | Pred. No. 0.13; | | |
| Matches 117; Conservative | 0; | Mismatches 131; | Indels 0; | Gaps 0; |

```

QY      375  AGCCCGTGGGCGGATCTCTCTCATTTGACTCTCTCCAAGATCACTCGAGAAAGTCAAAGCAC 434
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      6431  ACCCCGTGCGCTGCAACTCTCTTGGAGAAATCAAAGCTTTGCGCGAGGCCACGACGCTT 6490
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      435  TTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCC 494
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      6491  TCCGCTCCTCCCTCAGCTTGCCAGGCTGACTTCAACCAAGCTGGCCGAGCTGAACGCC 6550
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      495  AGCTTACCACCTTTGGGCACTTCAGCTCTCAACCGTATAACTTCAGCACTCTGGAAGACCTGA 554
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      6551  AGATCAAGAGCTTCCGCGTAGCCTCCAAACCCCTACACCTGGTTTACCATGAGAGCCCTGG 6610
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      555  ACACCAAGATGSAAGCTTCTGCAAGGTGGCCGTGAGAACCGAGTCAGGCAGCTGCATGAAG 614
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      6611  AGGAGACCTGGAGGAACCTTACAGAAATCATCAAGSAGAGGAGCTGGAGCTGCAGAAAG 6670
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      615  CCCACAGG 622
      |||||
Db      6671  AACAGCGG 6678

```

```

RESULT 14
US-09-668-313A-3
; Sequence 3, Application US/09668313A
; Patent No. 6503756
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Susan M. Freier
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF SYNTAXIN 4 INTERACTING PROTEIN EXPRESSION
; FILE REFERENCE: RTS-0127
; CURRENT APPLICATION NUMBER: US/09/668, 313A
; CURRENT FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 247
; SEQ ID NO 3
; LENGTH: 428
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(171)
US-09-668-313A-3

```

| | | | | |
|--------------------------|-------|------------------|-----------|-------------|
| Query Match | 2.9% | Score 38.2; | DB 4; | Length 428; |
| Best Local Similarity | 55.7% | Pred. No. 0.025; | | |
| Matches 73; Conservative | 0; | Mismatches 58; | Indels 0; | Gaps 0; |

[illegible]

RESULT 15
US-09-668-313A-17
; Sequence 17, Application US/09668313A
; Patent No. 6503756
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Susan M. Freiler
; APPLICANT: Jacqueline Wyatt

```

; TITLE OF INVENTION: ANTISENSE MODULATION OF SYNTAXIN 4 INTERACTING PROTEIN EXPRES
; FILE REFERENCE: RTS-0127
; CURRENT APPLICATION NUMBER: US/09/668,313A
; CURRENT FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 247
; SEQ ID NO 17
; LENGTH: 4439
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (300)...(1964)
US-09-668-313A-17

```

| | Query Match | 2.9%; | Score 38.2; | DB 4; | Length 4439; | |
|----|--|--------|-----------------|-------|--------------|-----------------------|
| | Best Local Similarity | 55.7%; | Pred. No. 0.11; | | | |
| | Matches | 73; | Conservative | 0; | Mismatches | 58; Indels 0; Gaps 0; |
| QY | 674 TGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA | | | | | 733 |
| Dh | 1806 TGGAGAAGCTTACACAGCAGATGGAATCAAGTACTTCATCAATCATGTAAACACAGACT | | | | | 1865 |
| QY | 734 ACTTGCTGGACCATCCCAAAATGACAGAGCTTACCAGTCTTTAGCTGACCTGAATAAT | | | | | 793 |
| Dh | 1866 ACATCCTGGATCCATCCCCTGATGAGTGTCTGAACTTATCTCGCTCAGAGAGGAATGAA | | | | | 1925 |
| QY | 794 CTCAGATTCTC | | | | | 804 |
| Dh | 1926 GAGGATTGCTC | | | | | 1936 |

Search completed: February 2, 2004, 06:15:57
Job time : 83.4247 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 11:04:35 ; Search time 4502.13 Seconds
(without alignments)
12961.661 Million cell updates/sec

Title: US-09-845-416-9_COPY_1000_3400

Perfect score: 2401

Sequence: 1 ggcagtcattcatgagagag.....tcggggcagtaacattgagc 2401

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST: *
1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estmu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_htc: *
9: gb_est1: *
10: gb_est2: *
11: gb_htc: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: em_gss_hum: *
18: em_gss_inv: *
19: em_gss_pln: *
20: em_gss_vrt: *
21: em_gss_fun: *
22: em_gss_mam: *
23: em_gss_mus: *
24: em_gss_pro: *
25: em_gss_rod: *
26: em_gss_phg: *
27: em_gss_vrl: *
28: gb_gss1: *
29: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 693.8 | 28.9 | 3870 | 11 BC036103 | BC036103 Homo sapi |
| 2 | 605.4 | 25.2 | 834 | 12 B1729851 | B1729851 603349511 |
| 3 | 565.8 | 23.6 | 3056 | 11 AK044536 | AK044536 Mus muscu |
| 4 | 557.4 | 23.2 | 3753 | 11 AK081426 | AK081426 Mus muscu |

| | | | | | |
|----|-------|------|------|-------------|--------------------|
| 5 | 557 | 23.2 | 824 | 9 AL556247 | AL556247 AL556247 |
| 6 | 555.4 | 23.1 | 801 | 14 CB991394 | CB991394 AGENCOURT |
| 7 | 555.4 | 23.1 | 1098 | 13 BX365572 | BX365572 BX365572 |
| 8 | 546 | 22.7 | 620 | 13 BQ640063 | BQ640063 he23904.Y |
| 9 | 520.2 | 21.7 | 770 | 10 BG719710 | BG719710 602690430 |
| 10 | 511 | 21.3 | 797 | 14 CB960722 | CB960722 AGENCOURT |
| 11 | 502.6 | 20.9 | 1298 | 11 AK087829 | AK087829 Mus muscu |
| 12 | 502.6 | 20.9 | 2135 | 11 AK013510 | AK013510 Mus muscu |
| 13 | 502.6 | 20.9 | 4437 | 11 AK036936 | AK036936 Mus muscu |
| 14 | 501 | 20.9 | 1384 | 11 AK075809 | AK075809 Mus muscu |
| 15 | 495.2 | 20.6 | 750 | 12 B1730168 | B1730168 603349711 |
| 16 | 491.6 | 20.5 | 704 | 10 BB610411 | BB610411 BB610411 |
| 17 | 486.6 | 20.3 | 579 | 9 AL121550 | AL121550 DKF2P762L |
| 18 | 485.8 | 20.2 | 652 | 10 BB629984 | BB629984 BB629984 |
| 19 | 485.8 | 20.2 | 728 | 14 CB228986 | CB228986 AGENCOURT |
| 20 | 458.4 | 19.1 | 854 | 9 AL196693 | AL196693 u153e10.Y |
| 21 | 455.8 | 19.0 | 591 | 2 HSM083521 | Bx490860 Homo sapi |
| 22 | 451.2 | 18.8 | 3051 | 11 BC036095 | BC036095 Homo sapi |
| 23 | 447 | 18.6 | 717 | 14 CB527785 | CB527785 UI-M-FY0- |
| 24 | 439 | 18.3 | 578 | 2 HSM075761 | Bx485574 Homo sapi |
| 25 | 436.6 | 18.2 | 665 | 14 BY742604 | BY742604 BY742604 |
| 26 | 424.2 | 17.7 | 663 | 12 BM488464 | BM488464 pgm2n.pk0 |
| 27 | 414.6 | 17.3 | 1490 | 11 BC009242 | BC009242 Homo sapi |
| 28 | 414 | 17.2 | 898 | 10 BF182065 | BF182065 601804604 |
| 29 | 410.4 | 17.1 | 599 | 10 BB666688 | BB666688 BB666688 |
| 30 | 409 | 17.0 | 845 | 14 CB177816 | CB177816 IS21C01.X |
| 31 | 408.8 | 17.0 | 593 | 14 CA988247 | CA988247 AGENCOURT |
| 32 | 407.2 | 17.0 | 554 | 14 CB613696 | CB613696 AMGNNUC:N |
| 33 | 404.4 | 16.8 | 493 | 14 CA888041 | CA888041 B0142C06- |
| 34 | 397.4 | 16.6 | 495 | 14 CA894775 | CA894775 B0187G06- |
| 35 | 392.2 | 16.3 | 650 | 14 BY714491 | BY714491 BY714491 |
| 36 | 389 | 16.2 | 644 | 13 BU313510 | BU313510 603540290 |
| 37 | 385.2 | 16.0 | 684 | 9 AL641565 | AL641565 AL641565 |
| 38 | 383.4 | 16.0 | 483 | 14 CA893902 | CA893902 B0182B01- |
| 39 | 381.8 | 15.9 | 423 | 9 AA460476 | AA460476 2x61e10.T |
| 40 | 380.4 | 15.8 | 633 | 9 AL796733 | AL796733 AL796733 |
| 41 | 375.4 | 15.6 | 514 | 9 AL871560 | AL871560 AL871560 |
| 42 | 368.4 | 15.3 | 794 | 13 BU424348 | BU424348 603234189 |
| 43 | 365.4 | 15.2 | 655 | 14 CA558919 | CA558919 K0251G03- |
| 44 | 349.6 | 14.6 | 546 | 4 BX516597 | Bx516597 RZPD Mus |
| 45 | 337.6 | 14.1 | 541 | 9 AL894729 | AL894729 AL894729 |

ALIGNMENTS

| | | | | | |
|------------|--|-------------|------|--------|-----------------|
| RESULT 1 | BC036103 | 3870 bp | mRNA | linear | HTC 04-MAR-2003 |
| LOCUS | BC036103 | | | | |
| DEFINITION | Homo sapiens, clone IMAGE:5274415, mRNA. | | | | |
| ACCESSION | BC036103 | | | | |
| VERSION | BC036103.1 | GI:23271310 | | | |
| KEYWORDS | HTC. | | | | |
| SOURCE | Homo sapiens (human) | | | | |
| ORGANISM | Homo sapiens | | | | |
| REFERENCE | 1 (bases 1 to 3870) | | | | |
| AUTHORS | Strausberg, R. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA | | | | |
| REMARK | NIH-MGC Project URL: http://mgc.nci.nih.gov | | | | |
| COMMENT | Contact: MGC help desk Email: cgabbs-remail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshayuki and Piero Carninci (RIKEN) CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 | | | | |

QY 808 GGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAGGTGAATTTGAAGCTCACACA 867
|||
Db 210 GGAGTCAGAGAGCTGATGAACCATGGCAAGATCTCCAGAGAGAAATTTGAAGCTCACACA 269
QY 868 GATGTTATTCACAACTTGATGAACACAGCCAAAATCTTGAGATCCCTGGAAGTTCC 927
|||
Db 270 GATATCTATCACAATCTTGATGAATAATGGCCAAAATCTTGAGATCCCTGGAAGTTCC 329
QY 928 GATGATGCACTCTGTTTACAAAGACGTTTGATTAACATGAATCTCAAGTGGAGTGAACCTT 987
|||
Db 330 GATGAAGACACCCCTGTTTACAAAGACGTTTGATTAACATGAATTTCAAGTGGAGTGAACCTT 389
QY 988 CGGAAAAAGTCTCTCAACATTAGGTCCCATTTTGAAGCCAGTTCTGACCAGTGAAGCGT 1047
|||
Db 390 CAGAAAAAGTCTCTCAACATTAGGTCCCATTTTGAAGCCAGTTCTGACCAGTGAAGCGT 449
QY 1048 CTGCACCTTTCTCTGACAGAACTTGTGTGCTACAGCTGAAGATGAATTAAGC 1107
|||
Db 450 TTGCATCTTTCTCTTTCAGAACTTCTGTTGCTACAGCTGAAGATGAATGAAGCTGAGC 509
QY 1108 CGGACGACACCTATTGAGGCGAATTTCAGACAGTTGACAGACAGAAAGATGATACATAGG 1167
|||
Db 510 CGTCAGGACACCCATCGTGTGATTTCCAGACAGTTCAAGACAGAAATGATATACATAGG 569
QY 1168 GCCTTCAAGAGGGAATTGAAAATAAGAACCTGTATCATGAGTACTCTTGAGACTGTA 1227
|||
Db 570 GCCTTCAAGAGGGAATTGAAAATAAGAACCTGTATCATGAGTACTCTTGAGACTGTA 629
QY 1228 CGAATATTTCTGACAGAGCAGCTTTTGAAGGACTAGAGAAACTC-TACCAGAGCCCGAG 1286
|||
Db 630 AGAATATTTCTGACAGAGCAGCTTTTGAAGGACTAGAGAAACTC-TACCAGAGCCCGAG 689
QY 1287 AGAGCTGCCTCTCTGAGGAGAGAGAGCCCAAGATGTCTACCTGGCTTTACGAAAGACGGCTGA 1346
|||
Db 690 AGAAGTGCCTCTCTGAGGAGAGAGAGCTCAGAATGTCTACCTGGCTTTACGAAAGACGGCTGA 749
QY 1347 GGAGTCAATACTGAGTGGGAAAAATTGAACCTGCATC-CGCTGACTGGCAGAGAAAA 1404
|||
Db 750 AGAGGTCAACGCTGAATGGGACCAATTGAACCTGCCTCAAGCTGATTGGCAGAGAAAA 808
RESULT 3 3056 bp mRNA linear HTC 05-DEC-2002
AK044536
LOCUS AK044536
DEFINITION Mus musculus adult retina cDNA, RIKEN full-length enriched library,
clone: A930019F21 Product: dystrophin, muscular dystrophy, full
insert sequence.
ACCESSION AK044536
VERSION AK044536.1 GI:26090404
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Carninci, P. and Hayashizaki, Y.
1 High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
2 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumi, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Saito, T., Okazaki, Y., Gotobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
Quackenbush, J., Schriml, L. M., Stabili, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bul, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seta, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Toyo-oka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S.
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851
5
The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3056)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
CDNA library was prepared and sequenced in Mouse Genome
Encyclopedia project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Retina RNA was provided by Dr. Stefano Gustincich (Department of
Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA
02115, USA) whose assistance is gratefully acknowledged. Please
visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.
Location/Qualifiers
1. 3056
/organism="Mus musculus"

FEATURES
source

| | | | | |
|-----------------------|-----------------|---|-----------|--------------|
| | | /mol_type="mRNA" | | |
| | | /strain="C57Bl/6J" | | |
| | | /db_xref="FANTOM,DB:A930019F21" | | |
| | | /db_xref="taxon:10090" | | |
| | | /clone="A930019F21" | | |
| | | /tissue_type="retina" | | |
| | | /clone_lib="RIKEN full-length enriched mouse cDNA library" | | |
| | | /dev_stage="adult" | | |
| misc_feature | | 1..3056 | | |
| | | /note="dystrophin, muscular dystrophy (MGD MGI:94909, GB NM_007868, evidence: BLASTN, 100%, match=999)" | | |
| BASE COUNT | 1072 a | 567 c | 607 g | 810 t |
| ORIGIN | | | | |
| Query Match | 23.6%; | Score 565.8; | DB 11; | Length 3056; |
| Best Local Similarity | 82.5%; | Pred. No. 2.7e-108; | | |
| Matches 648; | Conservative 0; | Mismatches 137; | Indels 0; | Gaps 0; |
| QY | 1 | GGCAGTTCATTGATGAGAGTGAAGTAAACCTGACCGTTATCAACAGCTTTAGAAGAA | 60 | |
| Db | 1044 | GACAGTTCATTGATGAGACGGAAGTAAATCTGATAGTTACCAAACTGCTTTAGAAGAA | 1103 | |
| QY | 61 | GTATTATCGTGGCTTCTTCTGCTGAGACACATTTGCAAGCACAAGAGAGATTTCTAAT | 120 | |
| Db | 1104 | GTACTTTCATGGCTTCTTCTGCCGAGGATACATTTGCCAGCACAGAGAGATTTCAAAAT | 1163 | |
| QY | 121 | GATGTGAAGTGTGTAAGACCAAGTTTCATCTCATGATGAGGGGTACATGATGATTTGACA | 180 | |
| Db | 1164 | GATGTGAAGAAGTGAAGAAGACAGTTTCATGCTCATGAGGGATTTCATGATGATCTGACA | 1223 | |
| QY | 181 | GCCCCATCAGGGCGGGTGGTATATTTCTACAATTGGGAAGTAAGCTGATTTGGAACAGGA | 240 | |
| Db | 1224 | TCTCATCAAGGACTTGTGGTATGTTCTACAGTTAGGAAGTCAACTAGTTGGAAGGCG | 1283 | |
| QY | 241 | AAATTATCAGAAGATGAAGAACTGAAGTACAAAGAGAGATGAATCTCTAAATTCAAGA | 300 | |
| Db | 1284 | AAATTATCAGAAGATGAAGAACTGAAGTGCAAAGAACAAATGAATCTCTAAATTCAAGA | 1343 | |
| QY | 301 | TGGGAATGCCTCAGGCTAGCTAGCATGGAACAAAGCAATTACATAGAGTTTAAATG | 360 | |
| Db | 1344 | TGGGAATGCTCTCAGGCTAGCTAGCATGGAACAAAGCAATTACACAAAGTTCTAATG | 1403 | |
| QY | 361 | GATCTCCAGATCAGAACTGAAGAAGTTGAATGACTGGCTTAACAAAAAGAGAAAGA | 420 | |
| Db | 1404 | GATCTCCAGATCAGAAATTAAGAAGTACATGACTGGTTAACAAAACTGAAGAGAGA | 1463 | |
| QY | 421 | ACAAGGAAAATGAGGAAGCCTCTTGACCTGATCTTGAAGACCTTAAAGCCCAAGTA | 480 | |
| Db | 1464 | ACTAAGAAAATGAGGAAGAGCCCTTTGACCTGATCTTGAAGATCTAAAAATGCCAAGTA | 1523 | |
| QY | 481 | CAACACATTAAGGTGCTTCAAGAGATCTAGAACAAGAACAAAGTCAGGGTCAATTCTCTC | 540 | |
| Db | 1524 | CAACACATTAAGGTGCTTCAAGAAGATCTAGAACAAGAGCAGGTCAAGGTCAACTCGCTC | 1583 | |
| QY | 541 | ACTCACATGGTGTGGTAGTTGATGAATCTAGTGAGATCAAGCAACTGCTGCTTTGGA | 600 | |
| Db | 1584 | ACTCACATGGTAGTAGTGGTGTGATGAATCCAGCGGTGATCATGCAACAGCTGCTTTGGA | 1643 | |
| QY | 601 | GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGACAGAGACCCG | 660 | |
| Db | 1644 | GAACAACCTTAAGGTACTGGGAGATCGATGGGCAATATCTGCAAGATGACTGAAGACCCG | 1703 | |
| QY | 661 | TGGGTTCTTTTACAAGACACTCATAGATTACTGCAACAGTTCCCTGAGCCTTGAAAAAG | 720 | |
| Db | 1704 | TGGATTGTTTACAAGATATTCTTCTAAATGGCAGCATTTTACTGAAGAACAGTGCCTT | 1763 | |
| QY | 721 | TTTCTTGCTGGCTTACAGAGCTGAACAACATGCTCCATGACAGATGCTACCCGT | 780 | |
| Db | 1764 | TTTATGATCATGGCTTTTCAGAAAAAGAGATGCAATGAAGAACATTCAAGCAAGTGGCTTT | 1823 | |
| QY | 781 | AAGA 785 | | |
| Db | 1824 | AAAGA 1828 | | |

| | |
|------------|---|
| RESULT 4 | |
| AK081426 | |
| LOCUS | |
| DEFINITION | AK081426 3753 bp mRNA linear HTC 05-DEC-2002 |
| | Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched |
| | library, clone:Cl30016K19 product:dystrophin related protein 2, |
| | full insert sequence. |
| ACCESSION | AK081426 |
| VERSION | AK081426.1 GI:26349154 |
| KEYWORDS | HTC; CAP trapper. |
| SOURCE | Mus musculus (house mouse) |
| ORGANISM | Mus musculus |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. |
| REFERENCE | |
| AUTHORS | 1 Carninci,P. and Hayashizaki,Y. |
| TITLE | High-efficiency full-length cDNA cloning |
| JOURNAL | Meth. Enzymol. 303, 19-44 (1999) |
| MEDLINE | 99279253 |
| PUBMED | 10349636 |
| REFERENCE | |
| AUTHORS | 2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., |
| | Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. |
| TITLE | Normalization and subtraction of cap-trapper-selected cDNAs to |
| JOURNAL | prepare full-length cDNA libraries for rapid discovery of new genes |
| MEDLINE | Genome Res. 10 (10), 1617-1630 (2000) |
| PUBMED | 20499374 |
| REFERENCE | |
| AUTHORS | 3 |
| | Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., |
| | Konno,H., Akiyama,J., Nishi,K., Katsunai,T., Tashiro,H., Itoh,M., |
| | Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., |
| | Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., |
| | Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., |
| | Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., |
| | Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. |
| TITLE | RIKEN integrated sequence analysis (RISA) system--384-format |
| JOURNAL | sequencing pipeline with 384 multicapillary sequencer |
| MEDLINE | Genome Res. 10 (11), 1757-1771 (2000) |
| PUBMED | 20530913 |
| REFERENCE | |
| AUTHORS | 4 |
| | Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., |
| | Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., |
| | Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., |
| | Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., |
| | Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., |
| | Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., |
| | Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G., |
| | Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M., |
| | Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., |
| | Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., |
| | Carninci,P., de Bonaudo,M.F., Brownstein,M.J., Bult,C., |
| | Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., |
| | Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., |
| | Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., |
| | Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., |
| | Sato,K., Schonbach,C., Seta,T., Shibata,Y., Storch,K.F., Suzuki,H., |
| | Toyo-oka,K., Wang,K.H., Weitz,C., Whitaker,C., Wilting,L., |
| | Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawai,H., Kohetsuki,S. |
| | and Hayashizaki,Y. |
| TITLE | Functional annotation of a full-length mouse cDNA collection |
| JOURNAL | Nature 409 (6821), 685-690 (2001) |
| MEDLINE | 21085660 |
| PUBMED | 11217851 |
| REFERENCE | |
| AUTHORS | 5 |
| | The FANTOM Consortium and the RIKEN Genome Exploration Research |
| TITLE | Group Phase I & II Team. |
| JOURNAL | Analysis of the mouse transcriptome based on functional annotation |
| REFERENCE | of 60,770 full-length cDNAs |
| | Nature 420, 563-573 (2002) |
| | 6 (bases 1 to 3753) |

| | | | |
|----|------|---|------|
| QY | 2043 | TGACCAGCCCCATGGATATCTCTGCAGATTATTAATTGTTTGACCACTATTATGACCGCCT | 2102 |
| Db | 1494 | TGAACATGTGATGGATGTGTGGAGGTCAATTCACCTGCTTGACTGCGCTGTATGAACGACT | 1553 |
| QY | 2103 | GGAGCAAGAGCACACAATTGGTCAACGCTCCCTCTCTGCGTGGATATGTGTGTAAGT | 2162 |
| Db | 1554 | GGAGGAGGAAAGAGCGATCCTGTGTCATGTGCCACTGTGTGTAGACATGAGCCTCAACTG | 1613 |
| QY | 2163 | GCTGCTGAATGTTATGATACGGGACGAACAGGAGGATCCGTCTCTGTTTAAAC | 2222 |
| Db | 1614 | GCTCCTCAATGTTTTTGTAGTGTGTGCAGTGGAAAGATGCGACATTGTCCCTTAAAGC | 1673 |
| QY | 2223 | TGGCATCATTTCCCTGTGTAAAGCACATTTTGAAGACAGTACAGATACCTTTCAAGCA | 2282 |
| Db | 1674 | TGGCATCGCATGCGCTGTGTGGCACCGGAAGTGAAGAAGAAAACCTTCAGTATCTCTTCAGCCA | 1733 |
| QY | 2283 | AGTGGCAAGTTCACACAGGATTTTGTGACCCAGCGCAGGCTGGGCTCCTTCTGCATGATTC | 2342 |
| Db | 1734 | AGTAGCCCAATTCAGGCGAGCGCAGTGTGATCAACGCCCATCTTGGGTGCCCTGCTTCATGAAGC | 1793 |
| QY | 2343 | TATCCAAATTTCCAAGACAGTGGGTGAAGTTGCATCCTTTGGGGGCGAGTAAACATTGAGC | 2401 |
| Db | 1794 | CATTCAAGTCCCCCGTCAGCTGGGTGAAGTGGCAGCATTGGGGGCGAGCAATGTGGAGC | 1852 |

RESULT 5

| | | | | | |
|------------|----------------------------------|-------------|------|-------------------|-----------------|
| LOCUS | AL556247 | 824 bp | mRNA | linear | EST 31-MAY-2003 |
| DEFINITION | AL556247 Homo sapiens HELA CELLS | | | COT 25-NORMALIZED | Homo sapiens |
| ACCESSION | CDNA clone CS0DK001YB17 | 5-PRIME, | mRNA | sequence. | |
| VERSION | AL556247.2 | GI:31278051 | | | |

| | |
|----------|----------------------|
| SOURCE | Homo sapiens (human) |
| ORGANISM | Homo sapiens |

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Craniata; Eutheria; Primates; Catarrhini; Hominoidea; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

| REFERENCE | 1 (bases 1 to 824) |
|-----------|---|
| AUTHORS | Li, W.B., Gruber, C., Jesse, J. and Polayes, D. |
| TITLE | Full-length cDNA libraries and normalization |

Unpublished
On Feb 15, 2001 this sequence replaced gi:12898746

Contact: Genoscope
Genoscope - Centre National de Sequencage

Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email : segregf@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1955.r
Contact : Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/InvitrogenCorporation> 1600
Faraday Avenue Genoscope sequence ID : CS0DX001CA09QPL.

FEATURES

source

```

source
1. .824
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODK001YB17"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
211 a 210 c 206 g 197 t.
BASE COUNT
ORIGIN

```

| | | | | |
|---------------------------|---------|---------------------|-----------|-------------|
| Query Match | 23.2%; | Score 557; | DB 9; | Length 824; |
| Best Local Similarity | 100.0%; | Pred. No. 1.6e-106; | | |
| Matches 557; Conservative | 0; | Mismatches 0; | Indels 0; | Gaps 0; |

1845 CCACGAGACTCAACAACTTGTCTGGACCATCCCAAATGACAGAGCTCTACCACTCTT 1904

| | | | | | |
|----|------|-----------------|---|-------------------------------|-----|
| Db | 195 | CCACGAGACTCAACA | CACTGTGCTGGACCA | TCCCAAAATGACAGAGCTCTACAGTCTTT | 254 |
| QY | 1905 | AGCTGACCTGAA | TATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT | 1964 | |
| Db | 255 | AGCTGACCTGAA | TATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT | 314 | |
| QY | 1965 | GCAGAAAGGCCCTT | TGCTTGATCTCTTGAGCCTGTCAGCTGCATGTGATGCTTTGGACCA | 2024 | |
| Db | 315 | GCAGAAAGGCCCTT | TGCTTGATCTCTTGAGCCTGTCAGCTGCATGTGATGCTTTGGACCA | 374 | |
| QY | 2025 | GCACAACCTCAAGCA | AAAAATGACCAGCCCATGATATCCTGCAGATTATTAATTGTTGAC | 2084 | |
| Db | 375 | GCACAACCTCAAGCA | AAAAATGACCAGCCCATGATATCCTGCAGATTATTAATTGTTGAC | 434 | |
| QY | 2085 | CACATTTATGAC | CCGCTGGAGCAAGACACAACAATTTGGTCAACGTCCTCTGCGT | 2144 | |
| Db | 435 | CACATTTATGAC | CCGCTGGAGCAAGACACAACAATTTGGTCAACGTCCTCTGCGT | 494 | |
| QY | 2145 | GGATATGTGTCTGAA | CTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGATCCG | 2204 | |
| Db | 495 | GGATATGTGTCTGAA | CTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGATCCG | 554 | |
| QY | 2205 | TGTCCTGTCTTTAAA | ACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAACAAGTA | 2264 | |
| Db | 555 | TGTCCTGTCTTTAAA | ACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAACAAGTA | 614 | |
| QY | 2265 | CAGATACCTTTTCA | GCAAGTGGCAAGTTCAACAGATTTTGTGACCAGCGCAGGCTGGG | 2324 | |
| Db | 615 | CAGATACCTTTTCA | GCAAGTGGCAAGTTCAACAGATTTTGTGACCAGCGCAGGCTGGG | 674 | |
| QY | 2325 | CCTCCTTCTGCAT | GATTCATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGG | 2384 | |
| Db | 675 | CCTCCTTCTGCAT | GATTCATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGG | 734 | |
| QY | 2385 | GGGCGATACATGAGC | 2401 | | |
| Db | 735 | GGGCGATACATGAGC | 751 | | |

RESULT 6

1394

| | | | | | |
|------------|--|--------|------|--------|-----------------|
| LOCUS | CB991394 | 801 bp | mRNA | linear | EST 01-MAY-2003 |
| DEFINITION | AGENCOURT_13627932 NIH_MGC_148 Homo sapiens cDNA clone IMAGE:30336570 5', mRNA sequence. | | | | |

| | |
|-----------|-------------|
| ACCESSION | CB991394 |
| VERSION | CB991394.1 |
| | GI:30285818 |

| | |
|----------|----------------------|
| KEYWORDS | EST. |
| SOURCE | Homo sapiens (human) |

| ORGANISM | |
|---|--|
| Homo sapiens | |
| Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. | |

REFERENCE 1 (bases 1 to 801)

ETHORS

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished

COMMENT

```

Email: cgapbs-romail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: NDAM360 row: d column: 19
High quality sequence stop: 621.
Location/Qualifiers
FEATURES
source
1..801

```

```
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"
```

```
/clone="IMAGE:30336570"
/tissue_type="pre-ecamplic placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC 148"
/note="Organ: placenta; Vector: pBluescriptR; Site_1:
alt-XhoI; Site_2: BamH; Library is oligo-dT primed and
directionally cloned using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert
size 2.3 kb and normalized to ROT 5. This is a primary
library enriched for full-length clones and constructed
using the Cap-trapper method (Carninci, in preparation).
Library constructed by M. Brownstein (NIMH/NHGRI,
National Institutes of Health). Note: this is a NIH_MGC
Library."
```

BASE COUNT 204 a 210 c 199 g 188 t

ORIGIN

Query Match 23.1%; Score 555.4; DB 14; Length 801;
Best Local Similarity 99.8%; Pred. No. 3.4e-106;
Matches 556; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1845 CCACGAGACTCAACAACACTTGGTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTT 1904
Db 140 CCACGAGACTCAACAACACTTGGTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTT 199
QY 1905 AGCTGACCTGAATATGTCAGATTTCTGAGCTTATAGAGCTGCCATGAAGCTCGAAGACT 1964
Db 200 AGCTGACCTGAATATGTCAGATTTCTGAGCTTATAGAGCTGCCATGAAGCTCGAAGACT 259
QY 1965 GCAGAAAGCCCTTTGCTTGATCTCTGAGCCCTGTCAGCTGCATGTGAGCTTGGACCA 2024
Db 260 GCAGAAAGCCCTTTGCTTGATCTCTGAGCCCTGTCAGCTGCATGTGAGCTTGGACCA 319
QY 2025 GCACAACCTCAAGCAAAATGACAGCCCATGATATCTGACAGATTATTAATTGTTGAC 2084
Db 320 GCACAACCTCAAGCAAAATGACAGCCCATGATATCTGACAGATTATTAATTGTTGAC 379
QY 2085 CACTATTATGACCGCGCTGAGCAAGACACAACAATTGGTCAACGTCCTCTGCGT 2144
Db 380 CACTATTATGACCGCGCTGAGCAAGACACAACAATTGGTCAACGTCCTCTGCGT 439
QY 2145 GGATATGTGCTGACTGGCTGCTGAATGTTATGATACGGGACGACAGAGGATCCG 2204
Db 440 GGATATGTGCTGACTGGCTGCTGAATGTTATGATACGGGACGACAGAGGATCCG 499
QY 2205 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGACACAAGTA 2264
Db 500 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGACACAAGTA 559
QY 2265 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACACAGATTTTGTGACCAGCGGAGCTGGG 2324
Db 560 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACACAGATTTTGTGACCAGCGGAGCTGGG 619
QY 2325 CCTCTCTTGCATGATTTCTATCCAAATTCGAAGACAGTTGGTGAAGTTGATCCTTTGG 2384
Db 620 CCTCTCTTGCATGATTTCTATCCAAATTCGAAGACAGTTGGTGAAGTTGATCCTTTGG 679
QY 2385 GGGCAGTAACATTGAGC 2401
Db 680 GGGCAGTAACATTGAGC 696
```

RESULT 7
BX365572
LOCUS BX365572 1098 bp mRNA linear EST 05-MAY-2003
DEFINITION BX365572 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
ACCESSION BX365572
VERSION BX365572
KEYWORDS BX365572.1 GI:30366927
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS 1 (bases 1 to 1098)
TITLE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1955.r
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSI1AK001ZB11QPI.

FEATURES
source 1..1098
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK001YB21"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 286 a 271 c 268 g 262 t 11 others

ORIGIN

Query Match 23.1%; Score 555.4; DB 13; Length 1098;
Best Local Similarity 99.8%; Pred. No. 3.6e-106;
Matches 556; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1845 CCACGAGACTCAACAACACTTGGTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTT 1904
Db 204 CCACGAGACTCAACAACACTTGGTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTT 263
QY 1905 AGCTGACCTGAATATGTCAGATTTCTGAGCTTATAGAGCTGCCATGAAGCTCGAAGACT 1964
Db 264 AGCTGACCTGAATATGTCAGATTTCTGAGCTTATAGAGCTGCCATGAAGCTCGAAGACT 323
QY 1965 GCAGAAAGCCCTTTGCTTGATCTCTGAGCCCTGTCAGCTGCATGTGAGCTTGGACCA 2024
Db 324 GCAGAAAGCCCTTTGCTTGATCTCTGAGCCCTGTCAGCTGCATGTGAGCTTGGACCA 383
QY 2025 GCACAACCTCAAGCAAAATGACAGCCCATGATATCTGACAGATTATTAATTGTTGAC 2084
Db 384 GCACAACCTCAAGCAAAATGACAGCCCATGATATCTGACAGATTATTAATTGTTGAC 443
QY 2085 CACTATTATGACCGCGCTGAGCAAGACACAACAATTGGTCAACGTCCTCTGCGT 2144
Db 444 CACTATTATGACCGCGCTGAGCAAGACACAACAATTGGTCAACGTCCTCTGCGT 503
QY 2145 GGATATGTGCTGAACTGGCTGCTGAATGTTATGATACGGGACGAACAGAGGATCCG 2204
Db 504 GGATATGTGCTGAACTGGCTGCTGAATGTTATGATACGGGACGAACAGAGGATCCG 563
QY 2205 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGACACAAGTA 2264
Db 564 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGACACAAGTA 623
QY 2265 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACACAGATTTTGTGACCAGCGGAGCTGGG 2324
Db 624 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACACAGATTTTGTGACCAGCGGAGCTGGG 683
QY 2325 CCTCTCTTGCATGATTTCTATCCAAATTCGAAGACAGTTGGTGAAGTTGATCCTTTGG 2384
Db 684 CCTCTCTTGCATGATTTCTATCCAAATTCGAAGACAGTTGGTGAAGTTGATCCTTTGG 743
QY 2385 GGGCAGTAACATTGAGC 2401
Db 744 GGGCAGTAACATTGAGC 760
```


RESULT 8
BQ640063
LOCUS
DEFINITION BQ640063 620 bp mRNA linear EST 15-JUL-2002
he23g04.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he
Homo sapiens cDNA clone he23g04 5', mRNA sequence.
ACCESSION BQ640063.1 GI:21764522
VERSION BQ640063
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 620)
AUTHORS Wistow,G., Bernstein,S.L., Wyatt,M.K., Ray,S., Behal,A., Touchman
,J.W., Bouffard,G., Smith,D. and Peterson,K.
Expressed sequence tag analysis of human retina for the NEIBank
Project: Retbindin, an abundant, novel retinal cDNA and alternative
splicing of other retina-preferred gene transcripts
Mol. Vis. 8 (4), (2002) In press
JOURNAL Contact: Wistow G
COMMENT Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 23 row: g column: 04
Seq primer: M13RP1 reverse primer (ABI).
FEATURES
source
1..620
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="he23g04"
/tissue_type="Retina"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Human Retina cDNA (Un-normalized, unamplified
): hd/he"
/note="Organ: Eye; Vector: pSPORT1; Neural retina tissue
was dissected from two 80 year old donors with no observed
eye disease. 100ug of total RNA was used for library
construction. A directionally cloned cDNA library in the
pSPORT1 vector (Life Technologies) was constructed at
Bioserve Biotechnology (Laurel MD) essentially following
the protocols of the Superscript Plasmid System full
details of which are contained in the manufacturer's
instruction manual (http://www.lifetech.com/). First
strand synthesis was carried out using a Not I
primer-adaptor [5'-pgactagttctagatcgacggcgccccc(T)15-3'
]. EST analysis was performed on the unamplified library
at the NIH Intramural Sequencing Center (NISC)."
BASE COUNT 165 a 148 c 144 g 163 t
ORIGIN
Query Match 22.7%; Score 546; DB 13; Length 620;
Best Local Similarity 100.0%; Pred.No. 3.1e-104;
Matches 546; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2036 AGCAAAATGACCAGCCCATGGATATCTCGAGATTATATTGTTGACCACTATTATG 2095
DB 181 AGCAAAATGACCAGCCCATGGATATCTCGAGATTATATTGTTGACCACTATTATG 240
QY 2096 ACCGCTTGAGCAGACAGACACAATTTGTCACAGTCCCTCTCTGCGTGAATATGTGC 2155
DB 241 ACCGCTTGAGCAGACAGACACAATTTGTCACAGTCCCTCTCTGCGTGAATATGTGC 300
QY 2156 TGAAGTGGCTGCTGTAATGTTTATGATACGGGACGACAGGAGATCCGTCTCTCTT 2215
DB 301 TGAAGTGGCTGCTGTAATGTTTATGATACGGGACGACAGGAGATCCGTCTCTCTT 360
QY 2216 TTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACATACCTTT 2275
DB 361 TTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACATACCTTT 420
QY 2276 TCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACGACGCGAGGCTGGCTCTCTGC 2335
DB 421 TCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACGACGCGAGGCTGGCTCTCTGC 480
QY 2336 ATGATCTATCCAAATTCACAGACAGTTGGGTGAAGTGCATCCTTTGGGGGAGTACA 2395
DB 481 ATGATCTATCCAAATTCACAGACAGTTGGGTGAAGTGCATCCTTTGGGGGAGTACA 540
QY 2396 TTGAGC 2401
DB 541 TTGAGC 546
RESULT 9
BG719710
LOCUS
DEFINITION BG719710 770 bp mRNA linear EST 08-MAY-2001
602690430F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4822807 5',
mRNA sequence.
ACCESSION BG719710.1 GI:13998897
VERSION BG719710
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 770)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LAM10731 row: 1 column: 08
High quality sequence stop: 767.
FEATURES
source
1..770
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4822807"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/note="Organ: testis; Vector: pBluescript (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTT-TTTT-TVN-3',
size-selected for average insert size 2.2 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI/NHGR1, National

BASE COUNT 189 a 199 c 198 g 184 t
ORIGIN

Institutes of Health). Note: this is a NIH_MGC Library."

Query Match 21.7%; Score 520.2; DB 10; Length 770;
Best Local Similarity 98.0%; Pred. No. 8.4e-99;
Matches 548; Conservative 0; Mismatches 8; Indels 3; Gaps 2;

```
OY 1845 CCACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAATGACAGAGCTCTACAGTCTTT 1904
    |||||||
Db 150 CCACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAATGACAGAGCTCTACAGTCTTT 209
OY 1905 AGCTGACCTGAATATATGTCAGATTCTCAGCTTATAGACTGCCATGAACTCCGAAGACT 1964
    |||||||
Db 210 AGCTGACCTGAATATATGTCAGATTCTCAGCTTATAGACTGCCATGAACTCCGAAGACT 269
OY 1965 GCAGAGAGCCCTTTGCTTGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCTTGAGCCA 2024
    |||||||
Db 270 GCAGAGAGCCCTTTGCTTGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCTTGAGCCA 329
OY 2025 GCACAA--CCTCAAGCAAAATGACCCAGCCCATGATATCCTGCAGATATATTAATTGTTG 2082
    |||||
Db 330 GCACATCCTCAAGTCAGATGACAGCCCATGATATCCTGCAGATATATTAATTGTTG 389
OY 2083 ACCACTATTTATGACCCGCTGAGCAAGACACACAATTGGTCAACGTCCTCTCTGC 2142
    |||||||
Db 390 ACCACTATTTATGACCCGCTGAGCAAGACACACAATTGGTCAACGTCCTCTCTGC 449
OY 2143 GTGATATGTGTCTGTAAGTGGCTGCTGATGTTTATGATACGGGACGACAGGAGGATC 2202
    |||||||
Db 450 GTGATATGTGTCTGTAAGTGGCTGCTGATGTTTATGATACGGGACGACAGGAGGATC 509
OY 2203 CGTGTCTGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTGGAGACAAG 2262
    |||||||
Db 510 CGTGTCTGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTGGAGACAAG 569
OY 2263 TACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGATTTTGTGACCAGCGCAGGCTG 2322
    |||||||
Db 570 TACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGATTTTGTGACCAGCGCAGGCTG 628
OY 2323 GGCTCTCTTCTGCATGATTTCTAATCCAAATTCCAAGACAGTTGGTGAAGTTCATCCTTT 2382
    |||||||
Db 629 GGCTCTCTTCTGCATGATTTCTAATCCAAATTCCAAGACAGTGGGGTGAAGTTCATCCTTT 688
OY 2383 GGGGCGAGTAACATTGAGC 2401
    |||||||
Db 689 GGGGCGAGTAACATTGAGC 707
```

RESULT 10
CB960722 797 bp mRNA linear EST 29-APR-2003
LOCUS
DEFINITION AGENCOURT_13761995 NIH_MGC_147 Homo sapiens cDNA clone
IMAGE:30344479 5', mRNA sequence.
CB960722
VERSION CB960722.1 GI:30216839
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: csapbs-r@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LINL at:
http://image.lnl.gov
Plate: NDM380 row: n column: 08
High quality sequence stop: 620.
Location/Qualifiers

FEATURES
source

```
1..797
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30344479"
/tissue_type="Human Placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_147"
/note="Organ: placenta; Vector: pBluescriptR; Site_1:
ali-XhoI; Site_2: BamH; Oligo-dT primed using primer
5'-TTTTTTTCTTTTCTTTTCTTTT-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NHGM/NHGRI, National Institutes of Health). Note: This is
a NIH_MGC library."
```

BASE COUNT 199 a 208 c 199 g 191 t
ORIGIN

Query Match 21.3%; Score 511; DB 14; Length 797;
Best Local Similarity 99.8%; Pred. No. 7.3e-97;
Matches 522; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
OY 1845 CCACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAATGACAGAGCTCTACAGTCTTT 1904
    |||||||
Db 168 CCACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAATGACAGAGCTCTACAGTCTTT 227
OY 1905 AGCTGACCTGAATATATGTCAGATTCTCAGCTTATAGACTGCCATGAACTCCGAAGACT 1964
    |||||||
Db 228 AGCTGACCTGAATATATGTCAGATTCTCAGCTTATAGACTGCCATGAACTCCGAAGACT 287
OY 1965 GCAGAGAGCCCTTTGCTTGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCTTGAGCCA 2024
    |||||||
Db 288 GCAGAGAGCCCTTTGCTTGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCTTGAGCCA 347
OY 2025 GCACAACTTCAAGCAAAATGACCCCATGATATCTGCAAGATTATTAATTGTTGAC 2084
    |||||||
Db 348 GCACAACTTCAAGCAAAATGACCCCATGATATCTGCAAGATTATTAATTGTTGAC 407
OY 2085 CACTATTTATGACCGCCTGAGCAAGACACACAATTGGTCAACGTCCTCTCTGCGT 2144
    |||||||
Db 408 CACTATTTATGACCGCCTGAGCAAGACACACAATTGGTCAACGTCCTCTCTGCGT 467
OY 2145 GGATATGTGTCTGAACTGGCTGCTGAATGTTATGATACGGGACGAACAGGAGATCCG 2204
    |||||||
Db 468 GGATATGTGTCTGAACTGGCTGCTGAATGTTATGATACGGGACGAACAGGAGATCCG 527
OY 2205 TGTCTGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTGGAGACAAGTA 2264
    |||||||
Db 528 TGTCTGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTGGAGACAAGTA 587
OY 2265 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGAGATTGTTGACCAGCGCAGGCTGGG 2324
    |||||||
Db 588 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGAGATTGTTGACCAGCGCAGGCTGGG 647
OY 2325 -CCTCTCTCTGCATGATTTCTAATCCAAATTCCAAGACAGTTGGG 2366
    |||||||
Db 648 CCTCTCTCTGCATGATTTCTAATCCAAATTCCAAGACAGTTGGG 690
```

RESULT 11
AK087829 1298 bp mRNA linear HTC 05-DEC-2002
LOCUS
DEFINITION AK087829 1298 bp mRNA linear HTC 05-DEC-2002
Full-length enriched library, adult female ovary cDNA, RIKEN
muscular dystrophy, full insert sequence.
ACCESSION AK087829

VERSION AK097829.1 GI:26352777
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, K., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE
4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, K., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staudli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boiffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyo-oka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohsaki, S. and Hayashizaki, Y.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409 (6821), 685-690 (2001)
MEDLINE 21085660
PUBMED 11217851
REFERENCE
5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
MEDLINE 11217851
PUBMED 12108560
REFERENCE
6 Adachi, J., Aizawa, K., Akimura, T., Aizawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,

| | | | |
|---------------------------|------|--|----------------------------------|
| TITLE | | Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. | |
| JOURNAL | | Direct Submission | |
| | | Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216) | |
| COMMENT | | cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/ URL: http://location/qualifiers | |
| FEATURES | | source | |
| | | 1..1298 | |
| | | /organism="Mus musculus" | |
| | | /mol_type="mRNA" | |
| | | /strain="C57BL/6J" | |
| | | /db_xref="FANTOM DB:E330026B12" | |
| | | /db_xref="taxon:10090" | |
| | | /clone="3330026B12" | |
| | | /sex="female" | |
| | | /tissue_type="ovary" | |
| | | /clone_lib="RIKEN full-length enriched mouse cDNA library" | |
| | | /dev_stage="2 days pregnant adult" | |
| | | 120..1142 | |
| | | /note="unnamed protein product; dystrophin, muscular dystrophy (MGD MG1:94909, GB NM_007868, evidence: BLASTN, 100%, match=999) | |
| | | putative" | |
| | | /codon_start=1 | |
| | | /protein_id="BAC40019.1" | |
| | | /db_xref="GI:26352778" | |
| | | /translation="MREHLKSHETQTCWDHPKMTLYQSLADLNVRFSAYRTAMKLRLOKALCDLDSLASACDLDQNDQMDLQINCLTTIYDRLEQHNHNLVNVPICVDMLNVLINVDYDGTGRIRYLSFKTGIIISLCKAHLEDKRYLFEQVASSGFLBQSRLLGLHDSIQIPROLGEVSFGSNIPEVSRSFQFANPKPEIEALFLDMRLBPQSMVLPVLHRAVAETAKHQAKCNICEKPIIGFRYRLSKHFNPDICQSCFESGRVAGHKHMYEMVEYCTPTTSGEDVRDPAVLKMKFRTRYRAKHPRMGYLPVQTVLEGDNMET" | |
| BASE COUNT | | 343 a | 332 c 296 g 327 t |
| ORIGIN | | | |
| Query Match | | 20.9%; | Score 502.6; DB 11; Length 1298; |
| Best Local Similarity | | 93.9%; | Pred. No. 4.5e-95; |
| Matches 523; Conservative | | 0; | Mismatches 34; Indels 0; Gaps 0; |
| QY | 1845 | CCAGGAGCTCAAACTTGTCTGGAGCCATCCAAATGACAGAGCTTACCACTTT | 1904 |
| DB | 1440 | CCAGGAGCCCAACCACTTGTGGAGCCACCCCAAAATGACAGAGCTTACCACTTT | 199 |
| QY | 1905 | AGCTGACCTGAATATGTCAGATTCTCAGCTTATAGGACTGCCATGAATCCGAAGACT | 1964 |
| DB | 200 | AGCTGACCTGAATATATGTCAGATTCTCCGCGTATAGGACTGCCATGAAGCTCAGAAAGCT | 259 |
| QY | 1965 | GCAGAAGCCCTTGTGCTGATCTCTGAGCCCTGTCAGCTGATGATGCTTGACCA | 2024 |
| DB | 260 | CCAGAAGCCCTTGTGCTGATCTCTGAGCCCTGTCAGCTGATGATGCTTGACCA | 319 |
| QY | 2025 | GCACAACCTCAAGCAAAATGACCCCATGATATCTGACATTAATTAATGTTGAC | 2084 |
| DB | 320 | GCACAACCTCAAGCAAAATGACCCCATGATATCTGACATTAATTAATGTTGAC | 379 |
| QY | 2085 | CAGTATTATGACCGCTGGAGCAAGAGACACAATTTGTCACGCTCCCTCTGCGT | 2144 |
| DB | 380 | TACATTTATGATCGCTGGAGCAAGAGACACAATCTGTCAATGTCCTCTGTGT | 439 |

QY 2145 GGATATGTGCTGCACTGGCTGCTGAATGTTATGATACGGACGAACAGGAGATCCG 2204
 |||||
 Db 440 GGATATGTGCTCAACTGGCTTCTCAATGTTATGATACGGACGAACAGGAGATCCG 499
 |||||
 QY 2205 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAGTA 2264
 |||||
 Db 500 TGTCTGTCTTTTAAACTGGCATCATTTCTGTGTAAGCACATTTGGAAGACAAGTA 559
 |||||
 QY 2265 CAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGACCGAGGAGCTGGG 2324
 |||||
 Db 560 CAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGACCGAGGAGCTGGG 619
 |||||
 QY 2325 CCTCCTTTCGATGATTCTATCCAAATTCGAAGACAGTTGGTGAAGTTGCATCCTTTGG 2384
 |||||
 Db 620 TCTTCTTTCGATGATTCTATCCAAATTCGAAGACAGTTGGTGAAGTTGCATCCTTTGG 679
 |||||
 QY 2385 GGGCAGTAACATGAGC 2401
 |||||
 Db 680 GGGCAGTAACATGAGC 696
 |||||
 RESULT 12
 AK013510
 LOCUS
 DEFINITION Mus musculus adult male hippocampus cDNA, RIKEN full-length enriched library, clone:2900010C03 product:dystrophin, muscular dystrophy, full insert sequence.
 AK013510
 AK013510.1 GI:12850899
 VERSION
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 AUTHORS
 1 Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE
 AUTHORS
 2
 3
 TITLE
 JOURNAL
 MEDLINE 20499374
 PUBMED 11042159
 REFERENCE
 AUTHORS
 1
 2
 3
 TITLE
 JOURNAL
 MEDLINE 20530913
 PUBMED 11076861
 REFERENCE
 AUTHORS
 4
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, K., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staudli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,

Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamuya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Nombarts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seta, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyo-oka, K., Wang, K. H., Weltz, C., Wittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohsaki, S. and Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 MEDLINE 21085660
 PUBMED 11217851
 5
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 5
 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 2135)
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hirakawa, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
 Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGAGTAAATTAATATCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI, 3' end: SstI.
 Host: SOLR.
 FEATURES
 source
 1. .2135
 /organism="Mus musculus"
 /mol_type="RNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM_DB:2900010C03"
 /db_xref="MGI:1905833"
 /db_xref="taxon:10090"
 /clone="2900010C03"
 /sex="male"
 /tissue_type="hippocampus"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 1. .2135
 /note="dystrophin, muscular dystrophy (MGD)[MGI:94909, GB|NM_007868, evidence: BLASTN, 100%, match=999]"
 /db_xref="MGI:94909"
 misc_feature
 563 a 457 c 438 g 677 t
 BASE COUNT
 ORIGIN

Query Match 20.9%; Score 502.6; DB 11; Length 2135;
Best Local Similarity 93.9%; Pred. No. 4.8e-95;
Matches 523; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 1845 CCACGAGCTCAAAACACTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTT 1904
DB 144 CCACGAGACCCAAACCACTTGTGGGACCAACCCAAATGACAGAGCTCTACAGTCTTT 203

QY 1905 AGCTGACCTGAATATGTGAGATTTCTGACCTTATAGAGCTGCCATGAAACTCCGAGACT 1964
DB 204 AGCTGACCTGAATATGTGAGATTTCTGCCGTATAGAGCTGCCATGAAAGCTCAGAGGCT 263

QY 1965 GCAGAAGGCCCTTTGCTTGATCTCTTGAGCCTGTGACGCTGCATGTGATGCCCTTGACCA 2024
DB 264 CCAGAAGGCCCTTTGCTTGATCTCTTGAGCCTGTGACGCTGCATGTGATGCCCTTGACCA 323

QY 2025 GCACACCTCAAGCAAAATGACCAAGCCCATGATATCCTGACATTTATTTATTTGTTGAC 2084
DB 324 GCACACCTCAAGCAAAATGACCAAGCCCATGATATCCTGACATTTATTTATTTGTTGAC 383

QY 2085 CACTATTATGACCGCCTGAGAGCAAGACACAACTTTGGTCAACGCTCCCTCTCTGCGT 2144
DB 384 TACATTTATGATCGCTGTGAGAGCAAGACACAACTTTGGTCAATGTTCCCTCTCTGTGT 443

QY 2145 GGATATGTGTCTGAACCTGGCTGCTGAATGTTATGATACGGGACGACAGAGGAGATCCG 2204
DB 444 GGATATGTGTCTGAACCTGGCTTCTCAATGTTATGATACGGGACGACAGAGGAGATCCG 503

QY 2205 TGTCTGTCTTTTAAACTGGCAATCTTCCCTGTGTAAGACACATTTGGAAGACAGTA 2264
DB 504 TGTCTGTCTTTTAAACTGGCAATCTTCTGTGTAAGACACATTTGGAAGACAGTA 563

QY 2265 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGAGATTTGTGACCGCAGCGCTGGG 2324
DB 564 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGCTGCTTTGTGACCGCTAGCGTGGG 623

QY 2325 CCTCCTTTCATGATTTCTATCCAAATTCACAGACAGTGGGTGAAGTTGCATCCTTTGG 2384
DB 624 TCTTCTTTCATGATTTCTATTCAAATCCCAAGACAGTGGGTGAAGTTGCATCCTTTGG 683

QY 2385 GGGCAGTACATTGAGC 2401
DB 684 GGGCAGTACATTGAGC 700

RESULT 13
AK036936 4437 bp mRNA linear HTC 05-DEC-2002
LOCUS
DEFINITION Mus musculus adult female vagina cDNA, RIKEN full-length enriched library, clone:9930028B14 product:dystrophin, muscular dystrophy, full insert sequence.
ACCESSION AK036936
VERSION AK036936.1 GI:26085504
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159

REFERENCE
1. .4437

AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861

REFERENCE
AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Stabli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamly, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombauts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyo-oka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohsaki, S. and Hayashizaki, Y.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409 (6821), 685-690 (2001)
MEDLINE 21085660
PUBMED 11217851

REFERENCE
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
MEDLINE 12127851
PUBMED 12127851

REFERENCE
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
location/Qualifiers
1. .4437

FEATURES
source


```
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:9930028B14"
/db_xref="taxon:10090"
/clone="9930028B14"
/sex="female"
/tissue_type="vagina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
29.1968
/note="dystrophin, muscular dystrophy (MGD MGI:94909,
GB|NM_007868, evidence: BLASTN, 100%, match=999)
putative"
polyA_signal 4417..4422
/note="putative"
polyA_site 4437
/note="putative"
BASE COUNT 1300 a 932 c 885 g 1320 t
ORIGIN
```

```
Query Match 20.9%; Score 502.6; DB 11; Length 4437;
Best Local Similarity 93.9%; Pred. No. 5.3e-95;
Matches 523; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
```

```
QY 1845 CCACGAGACTCAACAACTTGGTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTT 1904
Db 49 CCACGAGACCCCAACCACTGTTGGGACCAACCCCAAAATGACAGAGCTCTACCACTCTT 108
QY 1905 AGCTGACCTGAATATATGTCAGATTTCTCAGCTTATAGGACTGCCATGAAGCTCCGAAGACT 1964
Db 109 AGCTGACCTGAATATATGTCAGATTTCTCAGCTTATAGGACTGCCATGAAGCTCCGAAGGCT 168
QY 1965 GCAGAAAGCCCTTGTGCTGATCTCTTGAGCCCTGTCACTGTCATGTGATGCTTGGACCA 2024
Db 169 CCAGAAAGCCCTTGTGCTGATCTCTTGAGCCCTGTCACTGTCATGTGATGCTTGGACCA 228
QY 2025 GCACAACTCTCAAGCAAAATGACAGCCCATGATATCTGACAGATTATTAATTGTTGAC 2084
Db 229 GCACAACTCTCAAGCAAAATGACAGCCCATGATATCTGACAGATTATTAATTGTTGAC 288
QY 2085 CACTATTATGACCGCTGGAGCAAGACACAAATTGTCACAGTCCCTCTCTGCGT 2144
Db 289 TACAATTATGATCGTCTGGAGCAAGACACAAATTGTCACAGTCCCTCTCTGCGT 348
QY 2145 GGATATGTGTCTGAAGTGGCTGCTGATGTTATGATACGGAGCAAGAGGAGATCCG 2204
Db 349 GGATATGTGTCTGAAGTGGCTGCTGATGTTATGATACGGAGCAAGAGGAGATCCG 408
QY 2205 TGTCTGTCTTTAAAACTGGGATCATTTCCCTGTGTAAGACACATTTGAGAGACAAGTA 2264
Db 409 TGTCTGTCTTTAAAACTGGGATCATTTCTGTGTAAGACACATTTGAGAGACAAGTA 468
QY 2265 CAGATACCTTTTCAAGCAAGTGGCAAGTCAACAGAGATTTTGACCAGCGAGGCTGGG 2324
Db 469 CAGATACCTTTTCAAGCAAGTGGCAAGTCAACAGGCTTTTGACCAGCGTAGGCTGGG 528
QY 2325 CCTCCTTCTGCATGATTTATCCAAATTCAGAGACAGTTGGGTGAAGTTGCATCTTTGG 2384
Db 529 TCTTCTTCTGCATGATTTATTCAAATCCCAAGACAGTTGGGTGAAGTTGCTTCTTTGG 588
QY 2385 GGGCAGTAACATTGAGC 2401
Db 589 GGGCAGTAACATTGAGC 605
```

```
RESULT 14
LOCUS AK075809 1384 bp mRNA linear HTC 07-DEC-2002
DEFINITION Mus musculus 10 day old male pancreas cDNA, RIKEN full-length
dystrophin, full insert sequence.
ACCESSION AK075809
```

VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

AK075809.1 GI:26344655
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumi, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
11076861
4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
Quackenbush, J., Schiraldi, L.M., Stabli, P., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seta, T., Shibata, Y., Storch, K.F., Suzuki, H.,
Toyo-oka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohlsuki, S.
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1384)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,

TITLE
JOURNAL
Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A.,
Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers

FEATURES
source
1.1384

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:1810074E05"
/db_xref="MGI:1908112"
/db_xref="taxon:10090"
/clone="1810074E05"
/sex="male"
/tissue_type="pancreas"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 day old"
120.1142
/note="unamed protein product; dystrophin, muscular
dystrophy (MGD|MGI:94909, GB|NM_007868, evidence: BLASTN,
100%, match=999)
putative"
/codon_start=1
/protein_id="BAC35977.1"
/db_xref="GI:26344656"
/db_xref="MGI:94909"
/translation="MREHLKGETQTCDHPKMTLYQSLADLNVRFSAYRTAMKL
RLQKALCLDLISACDALDQNLKONDQPMDLQIINCLTIYDLQEHNNLVN
VPLQVDMCLNMLNLYDTRGTRIRVLSFKTGIIISLCKAHLEDKRYLFKQVASSTGR
CDQRLGLLHDSIQIPRLGEVASFGSNIPEVSRCFQFANNKPEIEALFLDMMR
LEPQSMVWLPLVLRHVAALTAHQAKNICKECPRIIGFRYRLSKHFNVDICQSCFSG
RVAKGHKMHYPMVEYCTPTTSGEDVRDPAKVLKNKFRTRKYPFAKHPRMGYLPVQTVLE
GDNMT"

BASE COUNT
ORIGIN
376 a 354 c 303 g 351 t

Query Match 20.9%; Score 501; DB 11; Length 1384;
Best Local Similarity 93.7%; Pred. No. 9.9e-95;
Matches 522; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

OY 1845 CCACGAGACTCAAAACACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTT 1904
DB 140 CCACGAGACCCCAAAACACTTGCTGGGACCAACCCCAAAATGACAGAGCTCTACAGTCTTT 199
OY 1905 AGCTGACCTGAATATGTCAGATTCTCAGCTTATAGACTGCCATGAACTCCGAAGACT 1964
DB 200 AGCTGACCTGAATATGTCAGATTCTCAGCTTATAGACTGCCATGAACTCCGAAGACT 259
OY 1965 GCAGAGAGCCCTTGTGCTTGGATCTCTGAGCCCTGTCAGCTGCATGTGATGCTTGGACCA 2024
DB 260 CCAGAGAGCCCTTGTGCTTGGATCTCTGAGCCCTGTCAGCTGCATGTGATGCTTGGACCA 319
OY 2025 GCACAACTCAAGCAAAATGACAGAGCCCATGATATCTGCAGATTATTAATGTTTGGAC 2084
DB 320 GCACAACTCAAGCAAAATGACAGAGCCCATGATATCTGCAGATTATTAATGTTTGGAC 379
OY 2085 CACTATTATGACCGCTGAGAGACAGACACAACAATTGTCACAGTCCCTCTCTGCGT 2144

DB 380 TACAATTATGATCGTCTGAGAGCAAGACACAACAATCTGTGTCATGTCCCTCTGTGT 439
OY 2145 GGATATGTGTCTGAACTGGCTGCTGTAATGTTATGATACGGGACGACAGAGGATCCG 2204
DB 440 GGATATGTGTCTGAACTGGCTGCTGTAATGTTATGATACGGGACGACAGAGGATCCG 499
OY 2205 TGTCTGTCTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAACAAGTA 2264
DB 500 TGTCTGTCTTTAAACTGGCATCATTTCTGTGTAAAGCACACTTTGGAACAAGTA 559
OY 2265 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTGTGACCAAGGCTGGG 2324
DB 560 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTGTGACCAAGGCTGGG 619
OY 2325 CTTCTCTGTCATGATTTCTATCCAAATTCACAGACAGTTGGGTGAAGTTCATCCTTTGG 2384
DB 620 TCTTCTCTGTCATGATTTCTATTCAAATTCACAGACAGTTGGGTGAAGTTCCTTTGG 679
OY 2385 GGGCAGTAACATTGAGC 2401
DB 680 GGGCAGTAACATTGAGC 696

RESULT 15
BI730168 750 bp mRNA linear EST 20-SEP-2001
LOCUS 603349711F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5357187 5',
DEFINITION mRNA sequence.
ACCESSION BI730168
VERSION BI730168.1 GI:15707181
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM11908 row: F column: 04
High quality sequence stop: 747.

FEATURES
source
1.750

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:5357187"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 94"

/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT
ORIGIN
240 a 161 c 178 g 170 t 1 others

Query Match 20.6%; Score 495.2; DB 12; Length 750;
Best Local Similarity 85.5%; Pred. No. 1.5e-93;
Matches 636; Conservative 0; Mismatches 93; Indels 15; Gaps 7;

OY 600 AGAACAACTTAAGGTATGGAGATCGATGGGCAACATCTGTAGATGACAGAGACCG 659

```
Db      13 AGAATATCATACTTCTTGGGGAACATTCACCTAAACATGTAAGTGAGCAAGAGCTG 72
QY      660 CTGGGTTCTTTTACAGACACTCATAGATTACTGCAACAGTTCCCCCTGAGACCTGGA AAA 719
Db      73 CTGGGAC-----AGAACTCATAGATTACTGCAAGAGTTCCCTCTGAGACCTGAGAA 125
QY      720 GTTCTTGCCCTGG-CTTACAGAAAGTGAAAC--AACTGCCAATGTCTCTACAGAGATGCTAC 776
Db      126 GTTCTTTCCTTGATTTACGGAAGCAGAAAACAGAGCTGCCAATGTCTCTACAGAGAGCTTC 185
QY      777 CCG-TAAGGAAAAGGCTCCTAGAGACTCCAAAGG--AGTAAAAGAGCTGATGAAACAATG 833
Db      186 CCGCTAAGGAGAGCTCCTAGCAAGACTCCAGGGGAGTCAAGAGAGCTGATGAAACCATG 245
QY      834 GCAAGACTCCCAAGGTGAAA-TTGAAGCTCACACAGATGTTTATCACAACCTGATGAAA 892
Db      246 GCAAGATCTCCAAGAGAGAAAGTTGAAACTCACACAGATATCTATCACAATCTTGATGAAA 305
QY      893 ACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAAGTCCGTGTACAAAGAC 952
Db      306 ATGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAAGTCCGTGTACAAAGAC 365
QY      953 GTTGGATTAACATGAACCTCAAGTGAAGTGAACCTTCGAAAAAAGTCTCTCAACATTAAGT 1012
Db      366 GTTGGATTAACATGAATTTCAAGTGAAGTGAACCTTCGAAAAAAGTCTCTCAACATTAAGT 425
QY      1013 CCCATTTGGAAGCCAGTTCTGACAGTGAAGCGTCTGCACTTTCTCTGACAGAACTTC 1072
Db      426 CCCATTTGGAAGCAAGTTCTGACAGTGAAGCGTTCGCACTTTCTCTCTGACAGAACTTC 485
QY      1073 TGGTGGGCTACAGCTGAAGATGATTAAGCCGCGCAAGCACTATTTGAGGCGACT 1132
Db      486 TTGTTGGCTACAGCTGAAGATGATGAAGTGAAGCGCGTCAAGCACTCGGTGATG 545
QY      1133 TTCCAGCAGTTCAGAAACA-GAACGATGTACATAGGGCTTCAAGAGGGAATGAAAACT 1191
Db      546 TCCCAAGCAGTTCAGAAACAAGATGATATACATAGGGCTTCAAGAGGGAATGAAAACT 605
QY      1192 AAAGAACTGTATCATGAGTACTTGTGAGACTGTACGAATATTTCTGACAGAGCAGCCT 1251
Db      606 AAAGAACTGTATCATGAGTACTTGTGAGACTGTGAGAAATATTTCTGACAGAGCAGCCT 665
QY      1252 TTGGAAGGACTAGAGAAACTCTACAGAGAGCCAGAGAGCTCCTCTGAGGAGAGAGCC 1311
Db      666 TTGGAAGGACTAGAGAAACTCTACAGAGAGCCAGAGAGACTGCTCCTGAAGAAAAGAGCT 725
QY      1312 CAGAAATGCACTCGGCTTCTACGA 1335
Db      726 CAGAAATGCACTCGGCTCCTACGA 749
```

Search completed: February 2, 2004, 06:03:12
Job time : 4507.46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 23:51:42 ; Search time 786.082 Seconds
(without alignments)
11131.886 Million cell updates/sec

Title: US-09-845-416-9_COPY_1000_3400

Perfect score: 2401

Sequence: 1 ggccagttcattgatggagag.....tgggggagtaacattgagc 2401

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2434939 seqs, 1822278265 residues

Total number of hits satisfying chosen parameters: 4869878

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------------------------------|
| 1 | 2401 | 100.0 | 3858 | 13 | US-09-845-416-9 Sequence 9, Appli |
| 2 | 2401 | 100.0 | 4825 | 13 | US-09-845-416-29 Sequence 29, Appli |
| 3 | 2401 | 100.0 | 4848 | 13 | US-09-845-416-35 Sequence 35, Appli |
| 4 | 2401 | 100.0 | 5060 | 13 | US-09-845-416-36 Sequence 36, Appli |
| 5 | 2250 | 93.7 | 3999 | 13 | US-09-845-416-6 Sequence 6, Appli |
| 6 | 2250 | 93.7 | 4966 | 13 | US-09-845-416-28 Sequence 28, Appli |
| 7 | 2250 | 93.7 | 4990 | 13 | US-09-845-416-34 Sequence 34, Appli |
| 8 | 2067 | 86.1 | 4182 | 13 | US-09-845-416-2 Sequence 2, Appli |
| 9 | 2067 | 86.1 | 5149 | 13 | US-09-845-416-27 Sequence 27, Appli |
| 10 | 1839.4 | 76.6 | 5462 | 13 | US-10-149-736-41 Sequence 41, Appli |
| 11 | 1737 | 72.3 | 3531 | 13 | US-09-845-416-10 Sequence 10, Appli |
| 12 | 1737 | 72.3 | 4498 | 13 | US-09-845-416-30 Sequence 30, Appli |
| 13 | 1735.6 | 72.3 | 8689 | 13 | US-10-149-736-42 Sequence 42, Appli |
| 14 | 1735.6 | 72.3 | 11058 | 13 | US-09-845-416-1 Sequence 1, Appli |
| 15 | 1735.6 | 72.3 | 11443 | 13 | US-10-149-736-44 Sequence 44, Appli |

| | | | | | |
|----|--------|------|-------|----|---------------------------------------|
| 16 | 1735.6 | 72.3 | 12057 | 13 | US-10-149-736-47 Sequence 47, Appli |
| 17 | 1735.6 | 72.3 | 13957 | 10 | US-09-782-378A-22 Sequence 22, Appli |
| 18 | 1735.6 | 72.3 | 13957 | 10 | US-09-880-107-2284 Sequence 2284, Ap |
| 19 | 1735.6 | 72.3 | 13957 | 13 | US-10-149-736-1 Sequence 1, Appli |
| 20 | 1735.6 | 72.3 | 14082 | 13 | US-10-341-434-108 Sequence 108, App |
| 21 | 1725 | 71.8 | 5339 | 13 | US-10-149-736-40 Sequence 40, Appli |
| 22 | 1723 | 71.8 | 2169 | 13 | US-09-845-416-4 Sequence 4, Appli |
| 23 | 1695 | 70.6 | 3510 | 13 | US-09-845-416-12 Sequence 12, Appli |
| 24 | 1695 | 70.6 | 4476 | 13 | US-09-845-416-31 Sequence 31, Appli |
| 25 | 1506.8 | 62.8 | 13815 | 13 | US-10-149-736-2 Sequence 2, Appli |
| 26 | 1419 | 59.1 | 5417 | 13 | US-10-149-736-39 Sequence 39, Appli |
| 27 | 1376 | 57.3 | 1821 | 13 | US-09-845-416-13 Sequence 13, Appli |
| 28 | 1263 | 52.6 | 4414 | 13 | US-09-845-416-32 Sequence 32, Appli |
| 29 | 1251 | 52.1 | 3446 | 13 | US-09-845-416-14 Sequence 14, Appli |
| 30 | 989 | 41.2 | 1434 | 13 | US-09-845-416-15 Sequence 15, Appli |
| 31 | 708.2 | 29.5 | 11096 | 13 | US-10-149-736-4 Sequence 4, Appli |
| 32 | 692.2 | 28.8 | 1991 | 13 | US-09-845-416-3 Sequence 3, Appli |
| 33 | 690.6 | 28.8 | 10302 | 10 | US-09-782-378A-23 Sequence 23, Appli |
| 34 | 690.6 | 28.8 | 10302 | 13 | US-10-149-736-3 Sequence 3, Appli |
| 35 | 681.8 | 28.4 | 16531 | 13 | US-10-101-510-667 Sequence 667, App |
| 36 | 677 | 28.2 | 1667 | 13 | US-09-845-416-7 Sequence 7, Appli |
| 37 | 445 | 18.5 | 887 | 13 | US-10-149-736-35 Sequence 35, Appli |
| 38 | 387 | 16.1 | 387 | 13 | US-10-149-736-32 Sequence 32, Appli |
| 39 | 350 | 14.6 | 1340 | 13 | US-09-845-416-11 Sequence 11, Appli |
| 40 | 348 | 14.5 | 348 | 13 | US-10-149-736-31 Sequence 31, Appli |
| 41 | 331 | 13.8 | 331 | 13 | US-10-149-736-9 Sequence 9, Appli |
| 42 | 327 | 13.6 | 327 | 13 | US-10-149-736-8 Sequence 8, Appli |
| 43 | 324 | 13.5 | 324 | 13 | US-10-149-736-33 Sequence 33, Appli |
| 44 | 216 | 9.0 | 216 | 13 | US-10-149-736-34 Sequence 34, Appli |
| 45 | 160.2 | 6.7 | 256 | 9 | US-09-864-761-21956 Sequence 21956, A |

ALIGNMENTS

RESULT 1
US-09-845-416-9
; Sequence 9, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3858
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-9

| Query Match | Best Local Similarity | 100.0%; | Score 2401; | DB 13; | Length 3858; |
|---------------|-----------------------|---|-------------|--------|--------------|
| Matches 2401; | Conservative | 0; | Mismatches | 0; | Indels |
| QY | 1 | GGCAGTTCATGTGAGAGAGTAACCTGACCCGTTATCAACAGCTTTAGAGAA | 60 | | |
| DB | 1000 | GGCAGTTCATGTGAGAGAGTAACCTGACCCGTTATCAACAGCTTTAGAGAA | 1059 | | |
| QY | 61 | GATATTCGTGCTTCTTCTGTCAGGACACATTTGCAAGCAAGAGAGATTCTAAT | 120 | | |
| DB | 1060 | GATATTCGTGCTTCTTCTGTCAGGACACATTTGCAAGCAAGAGAGATTCTAAT | 1119 | | |
| QY | 121 | GATGTGGAAGTGTGAAGAGAGCTTCTACTCATGAGGGGTACATGATGATTGACA | 180 | | |
| DB | 1120 | GATGTGGAAGTGTGAAGAGAGCTTCTACTCATGAGGGGTACATGATGATTGACA | 1179 | | |
| QY | 181 | GGCCATCAGGGCCGGGTTGGTAATTTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA | 240 | | |

Db 1180 GCCCATCAGGCGCGGTTGGTAATATTCTACAATTGGGAGTAAGCTGATTGGAACAGGA 1239
QY 241 AAATTATCAGAGATGAAGAACTGAAATAACAAGACAGATGAATCTCTAAATTCAAGA 300
Db 1240 AAATTATCAGAGATGAAGAACTGAAATAACAAGACAGATGAATCTCTAAATTCAAGA 1299
QY 301 TGGGAATGCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAATG 360
Db 1300 TGGGAATGCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAATG 1359
QY 361 GATCTCCAGATCAGAAACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAGAAAGA 420
Db 1360 GATCTCCAGATCAGAAACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAGAAAGA 1419
QY 421 ACAAGAAAAATGGAGGAAGAGCCCTCTTGAACCTGATCTTGAGACCTAAACGCCAAGTA 480
Db 1420 ACAAGAAAAATGGAGGAAGAGCCCTCTTGAACCTGATCTTGAGACCTAAACGCCAAGTA 1479
QY 481 CAACAACATAAGGTGCTTCAAGAAAGATCTAGAACAAAGAACAGTCAAGGTCAATTCTCTC 540
Db 1480 CAACAACATAAGGTGCTTCAAGAAAGATCTAGAACAAAGAACAGTCAAGGTCAATTCTCTC 1539
QY 541 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGAGATCAAGCAACTGCTGCTTTGGAA 600
Db 1540 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGAGATCAAGCAACTGCTGCTTTGGAA 1599
QY 601 GAACAACATAAGGTATTGGAGATCGATGGGCAACATCTGTAGATGGAACAGAACGCCG 660
Db 1600 GAACAACATAAGGTATTGGAGATCGATGGGCAACATCTGTAGATGGAACAGAACGCCG 1659
QY 661 TGGGTTCTTTTACAAGACACTCATAGATTACTGCAACAGTCCCCCTGACCCTGMAAAG 720
Db 1660 TGGGTTCTTTTACAAGACACTCATAGATTACTGCAACAGTCCCCCTGACCCTGMAAAG 1719
QY 721 TTTCTTGCCCTGGCTTACAGAAAGCTGAAACAACCTGCCAATGTCTTACAGAGATGCTACCCGT 780
Db 1720 TTTCTTGCCCTGGCTTACAGAAAGCTGAAACAACCTGCCAATGTCTTACAGAGATGCTACCCGT 1779
QY 781 AAGGAAAGGCTCTAGAAAGACTCCAAAGGAGTAAAGAGCTGATGAAACAATGGCAAGAC 840
Db 1780 AAGGAAAGGCTCTAGAAAGACTCCAAAGGAGTAAAGAGCTGATGAAACAATGGCAAGAC 1839
QY 841 CTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGMAAAACAGCCAA 900
Db 1840 CTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGMAAAACAGCCAA 1899
QY 901 AAAATCCTGAGATCCCTGGAAGGTTCCGATGATGAGTCTGTGTTACAAGAAGCTTTGGAT 960
Db 1900 AAAATCCTGAGATCCCTGGAAGGTTCCGATGATGAGTCTGTGTTACAAGAAGCTTTGGAT 1959
QY 961 AACATGAACCTTCAAGTGAAGTGAACCTTGGAAAAAGTCTCTCAACATTAGTCCCAATTG 1020
Db 1960 AACATGAACCTTCAAGTGAAGTGAACCTTGGAAAAAGTCTCTCAACATTAGTCCCAATTG 2019
QY 1021 GAAGCCAGTTCTGACCAGTGAAGGCTCTGCACTTTCTCTGCAAGAACTTCTGTGTGG 1080
Db 2020 GAAGCCAGTTCTGACCAGTGAAGGCTCTGCACTTTCTCTGCAAGAACTTCTGTGTGG 2079
QY 1081 CTACAGCTGAAGATGATGAATTAAGCCGGCAGCAGCACTTATGGAAGCGACTTTCCAGCA 1140
Db 2080 CTACAGCTGAAGATGATGAATTAAGCCGGCAGCAGCACTTATGGAAGCGACTTTCCAGCA 2139
QY 1141 GTTCAGAGCAGAACGATGTACATAGGSCCTTCAAGAGGGAATTGAAAACTAAAGAACTT 1200
Db 2140 GTTCAGAGCAGAACGATGTACATAGGSCCTTCAAGAGGGAATTGAAAACTAAAGAACTT 2199
QY 1201 GTATATCATGATACTCTTGAGACTGTACGAATATTTCTGACAGAGCAAGCCTTTGAAGGA 1260
Db 2200 GTATATCATGATACTCTTGAGACTGTACGAATATTTCTGACAGAGCAAGCCTTTGAAGGA 2259
QY 1261 CTAGAGAACTCTACCGAGAGCCCAAGAGCTGCCTCCTGAGAGAGAGCCCAAGAAATGTC 1320

Db 2260 CTAGAGAACTCTACCGAGAGCCCAAGAGAGCTGCCTCCTGAGGAGAGAGCCCAAGATGTC 2319
QY 1321 ACTGGCTTTCTACGAAGACAGGCTGAGAGGTCAATATACTGATGGAAAAATTGAACCTG 1380
Db 2320 ACTGGCTTTCTACGAAGACAGGCTGAGAGGTCAATATACTGATGGAAAAATTGAACCTG 2379
QY 1381 CACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAGAAGCTCCAGGAACCTCAA 1440
Db 2380 CACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAGAAGCTCCAGGAACCTCAA 2439
QY 1441 GAGGCCAGGATGAGCTGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGATCCTGG 1500
Db 2440 GAGGCCAGGATGAGCTGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGATCCTGG 2499
QY 1501 CAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATGACCTCGAGAAGTCAAGGCA 1560
Db 2500 CAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATGACCTCGAGAAGTCAAGGCA 2559
QY 1561 CTTGAGAGAAATTGCGCCTCTGAAGAAGACGTGAGCCACGTCAATGACCTTGCTGCG 1620
Db 2560 CTTGAGAGAAATTGCGCCTCTGAAGAAGACGTGAGCCACGTCAATGACCTTGCTGCG 2619
QY 1621 CAGCTTACCACTTTGGGCATTTGAGCTCTCAACCGTATTAACCTCAGCACTCTGGAAGACTG 1680
Db 2620 CAGCTTACCACTTTGGGCATTTGAGCTCTCAACCGTATTAACCTCAGCACTCTGGAAGACTG 2679
QY 1681 AACACCAGATGGAAGCTTCTGCAAGTGGCCGTGAGGACCGAGTCAAGGCAAGCTGCATGAA 1740
Db 2680 AACACCAGATGGAAGCTTCTGCAAGTGGCCGTGAGGACCGAGTCAAGGCAAGCTGCATGAA 2739
QY 1741 GCCCACAAGGACCTTTGGTCCAGATCTCAGACACTTTCTTTCCAAGTCTGTCCAGGGTCCC 1800
Db 2740 GCCCACAAGGACCTTTGGTCCAGATCTCAGACACTTTCTTTCCAAGTCTGTCCAGGGTCCC 2799
QY 1801 TGGGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCAAGAGACTCAAAACA 1860
Db 2800 TGGGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCAAGAGACTCAAAACA 2859
QY 1861 ACTTGCTGGAGCCATCCCAAAAATGACAGAGCTCTACCAAGCTTTAGCTGACCTGAATAAT 1920
Db 2860 ACTTGCTGGAGCCATCCCAAAAATGACAGAGCTCTACCAAGCTTTAGCTGACCTGAATAAT 2919
QY 1921 GTGAGATCTCAGCTTATAGAGCTGCCATGAAACTCCGAAGACTGAGAAAGGCCCTTGC 1980
Db 2920 GTGAGATCTCAGCTTATAGAGCTGCCATGAAACTCCGAAGACTGAGAAAGGCCCTTGC 2979
QY 1981 TTGATCTCTTGAGCCTGTGAGCTGATGTGATGCTTGGACCAAGCAACTCAAGCAA 2040
Db 2980 TTGATCTCTTGAGCCTGTGAGCTGATGTGATGCTTGGACCAAGCAACTCAAGCAA 3039
QY 2041 AATGACAGCCCATGATATCTCGACAGATTAATTTGTAACCACTATTATGACCCG 2100
Db 3040 AATGACAGCCCATGATATCTCGACAGATTAATTTGTAACCACTATTATGACCCG 3099
QY 2101 CTGAGCAGAGACACACAATTTGGTCAACGTCCCTCTGCGTGATATGTGTGAAC 2160
Db 3100 CTGAGCAGAGACACACAATTTGGTCAACGTCCCTCTGCGTGATATGTGTGAAC 3159
QY 2161 TGGCTGTGAATGTTTATGATACCGGAGCAAGACAGGAGATCCGTGTCTCTTTAAA 2220
Db 3160 TGGCTGTGAATGTTTATGATACCGGAGCAAGACAGGAGATCCGTGTCTCTTTAAA 3219
QY 2221 ACTGGCATATTTCCCTGTGTAAAGCACATTTGGAAGACAGATACCATCTTTCAAAG 2280
Db 3220 ACTGGCATATTTCCCTGTGTAAAGCACATTTGGAAGACAGATACCATCTTTCAAAG 3279
QY 2281 CAAGTGGCAAGTTCACAAGGATTTTGTAACCAAGCGCAGGCTGGGCTCTCTGCAATGAT 2340
Db 3280 CAAGTGGCAAGTTCACAAGGATTTTGTAACCAAGCGCAGGCTGGGCTCTCTGCAATGAT 3339
QY 2341 TCTATCCAAATTTCCAAGACAGTTGGGTGAAGTGCATCTTTGGGGGCAAGTAAACATTAAG 2400
Db 3340 TCTATCCAAATTTCCAAGACAGTTGGGTGAAGTGCATCTTTGGGGGCAAGTAAACATTAAG 3399

QY 2401 C 2401
Db 3400 C 3400

RESULT 2
US-09-845-416-29
; Sequence 29, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 4825
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-29

Query Match 100.0%; Score 2401; DB 13; Length 4825;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCAGTTCATTGATGAGAGTGAAGTAACTGACCCGTTATCAACAGCTTTAGAGAA 60
Db 1757 GGCAGTTCATTGATGAGAGTGAAGTAACTGACCCGTTATCAACAGCTTTAGAGAA 1816
QY 61 GTATTATCGTGCTCTTTCTGCTGAGACACATGCAAGCACAAAGAGAGATTCTAAT 120
Db 1817 GTATTATCGTGCTCTTTCTGCTGAGACACATGCAAGCACAAAGAGAGATTCTAAT 1876
QY 121 GATGTGAAGTGTGAAGACCAAGTTTCATCTCATGAGGGTACATGATGATTGACA 180
Db 1877 GATGTGAAGTGTGAAGACCAAGTTTCATCTCATGAGGGTACATGATGATTGACA 1936
QY 181 GCCCATCAGGCGCGGTTGTAATATTTACAAATGGGAAGTAAGCTGATTGAACAGGA 240
Db 1937 GCCCATCAGGCGCGGTTGTAATATTTACAAATGGGAAGTAAGCTGATTGAACAGGA 1996
QY 241 AAATATCAGAAAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTAATTCAGA 300
Db 1997 AAATATCAGAAAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTAATTCAGA 2056
QY 301 TGGGAATGCTCAGGGTAGCTAGCATGGAAGAAACAAAGCAATTTACATAGATTTAATG 360
Db 2057 TGGGAATGCTCAGGGTAGCTAGCATGGAAGAAACAAAGCAATTTACATAGATTTAATG 2116
QY 361 GATCTCCAGAAATCAGAAACTGAAGAAGTTGAATGACTGGCTAACAAAAACAGAGAAAGA 420
Db 2117 GATCTCCAGAAATCAGAAACTGAAGAAGTTGAATGACTGGCTAACAAAAACAGAGAAAGA 2176
QY 421 ACAAGAAATGAGAGAGAGCCCTCTGGACCTGATCTTGAAAGACCTTAAACGCCAAGTA 480
Db 2177 ACAAGAAATGAGAGAGAGCCCTCTGGACCTGATCTTGAAAGACCTTAAACGCCAAGTA 2236
QY 481 CAACAACATAAGGTGCTTCAAGAGATCTAGAACAAAGCAAGTCAGGGTCAATTCTCTC 540
Db 2237 CAACAACATAAGGTGCTTCAAGAGATCTAGAACAAAGCAAGTCAGGGTCAATTCTCTC 2296
QY 541 ACTCACATGTTGGTGTAGTGTGATGAATCTAGTGAAGATCAGCAACTGCTGCTTTGAA 600
Db 2297 ACTCACATGTTGGTGTAGTGTGATGAATCTAGTGAAGATCAGCAACTGCTGCTTTGAA 2356
QY 601 GAACAACCTTAAGGTATTTGGAGATCGATGGGCAACATCTGTAGATGACAGAGACCGC 660

Db 2357 GAACAACCTTAAGGTATTTGGAGATCGATGGGCAACATCTGTAGATGACAGAGACCGC 2416
QY 661 TGGGTTCTTTTACAGACACTCATAGATTACTGCAACAGTTTCCCTCGACCTGGAAAAAG 720
Db 2417 TGGGTTCTTTTACAGACACTCATAGATTACTGCAACAGTTTCCCTCGACCTGGAAAAAG 2476
QY 721 TTTCTTGCTGCTTACAGAAAGCTGAACAACTGCCAATGTCTCTACAGATGCTACCCGT 780
Db 2477 TTTCTTGCTGCTTACAGAAAGCTGAACAACTGCCAATGTCTCTACAGATGCTACCCGT 2536
QY 781 AAGGAAAGGCTCTAGAGACTCCAAAGGAGTAAAGAGCTGATGAACATGCGCAAGAC 840
Db 2537 AAGGAAAGGCTCTAGAGACTCCAAAGGAGTAAAGAGCTGATGAACATGCGCAAGAC 2596
QY 841 CTCCAAGTGAAATGAACTCACAAGATGTTTATCAACAACCTGGATGAAACAGCCAA 900
Db 2597 CTCCAAGTGAAATGAACTCACAAGATGTTTATCAACAACCTGGATGAAACAGCCAA 2656
QY 901 AAATCCTGAGATCCCTGGAAGGTTCCGATGATGCACTCCTGTACAAAGACGTTGGAT 960
Db 2657 AAATCCTGAGATCCCTGGAAGGTTCCGATGATGCACTCCTGTACAAAGACGTTGGAT 2716
QY 961 AACATGAACCTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1020
Db 2717 AACATGAACCTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2776
QY 1021 GAAGCCAGTTCTGACAGTGAAGCGTCTGACCTTCTCTGACAGAACTCTGTGTGG 1080
Db 2777 GAAGCCAGTTCTGACAGTGAAGCGTCTGACCTTCTCTGACAGAACTCTGTGTGG 2836
QY 1081 CTACAGCTGAAAGATGATGAATTAAGCCGGCAGCACTATTGAGGCGAATTCCAGCA 1140
Db 2837 CTACAGCTGAAAGATGATGAATTAAGCCGGCAGCACTATTGAGGCGAATTCCAGCA 2896
QY 1141 GTTCAGAAAGCAGACGATGTACATAGGGCTTCAAGAGGGAATTGAACCTAAGAACT 1200
Db 2897 GTTCAGAAAGCAGACGATGTACATAGGGCTTCAAGAGGGAATTGAACCTAAGAACT 2956
QY 1201 GTAATCATGATGATCTTGAAGACTGTACGAATATTTCTGACAGAGACGCTTTGGAAGGA 1260
Db 2957 GTAATCATGATGATCTTGAAGACTGTACGAATATTTCTGACAGAGACGCTTTGGAAGGA 3016
QY 1261 CTAGAGAACTCTACAGAGAGAGCCAGAGAGCTGCTCTGAGGAGAGAGCCCAAGATGTC 1320
Db 3017 CTAGAGAACTCTACAGAGAGAGCCAGAGAGCTGCTCTGAGGAGAGAGCCCAAGATGTC 3076
QY 1321 ACTCGGCTTCTACGAAAGCAGGCTGAGAGGCTCAATCTGAGTGGGAAAAATTGAACCTG 1380
Db 3077 ACTCGGCTTCTACGAAAGCAGGCTGAGAGGCTCAATCTGAGTGGGAAAAATTGAACCTG 3136
QY 1381 CACTCCGCTGACTGCGAGAGAAAAATAGATGAGACCCCTGAAAGACTCCAGAACTTCAA 1440
Db 3137 CACTCCGCTGACTGCGAGAGAAAAATAGATGAGACCCCTGAAAGACTCCAGAACTTCAA 3196
QY 1441 GAGGCCAGGATGAGTGAACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGATCCTGG 1500
Db 3197 GAGGCCAGGATGAGTGAACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGATCCTGG 3256
QY 1501 CAGCCCGTGGGCGATCTCTCATTTGACTCTCTCAAGATCACTCAAGTCAAGGCA 1560
Db 3257 CAGCCCGTGGGCGATCTCTCATTTGACTCTCTCAAGATCACTCAAGTCAAGGCA 3316
QY 1561 CTTGAGAGAGAAATGCGCCTCTGAAGAGAAAGTGAAGCCAGCTGATGACCTTGTCTGC 1620
Db 3317 CTTGAGAGAGAAATGCGCCTCTGAAGAGAAAGTGAAGCCAGCTGATGACCTTGTCTGC 3376
QY 1621 CAGCTTACCACTTTGGGCAATTCAGCTCTCAGCGTATAACCTCAGCACTCTGGAAGACCTG 1680
Db 3377 CAGCTTACCACTTTGGGCAATTCAGCTCTCAGCGTATAACCTCAGCACTCTGGAAGACCTG 3436
QY 1681 AACACCAAGATGAAGCTTCTGCAAGTGGCCCTGAGAGACCGAGTCAAGCACTGCATGAA 1740
Db 3437 AACACCAAGATGAAGCTTCTGCAAGTGGCCCTGAGAGACCGAGTCAAGCACTGCATGAA 3496

QY 1741 GCCCAGAGGACTTTGGTCCAGCATCTCAGCATTCTTTCCACGCTGTCCAGGGTCCC 1800
| | | | |
Db 3497 GCCCAGAGGACTTTGGTCCAGCATCTCAGCATTCTTTCCACGCTGTCCAGGGTCCC 3556
| | | | |
QY 1801 TGGGAGAGAGCCATCTCGCCAAACAAGTGGCCCTACTATATCAACCACGAGACTCAACA 1860
| | | | |
Db 3557 TGGGAGAGAGCCATCTCGCCAAACAAGTGGCCCTACTATATCAACCACGAGACTCAACA 3616
| | | | |
QY 1861 ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAT 1920
| | | | |
Db 3617 ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAT 3676
| | | | |
QY 1921 GTCAGATTCTCAGCTTATAGAGCTGCCATGAACTCCGAAGACTGCAGAGGCCCTTTGC 1980
| | | | |
Db 3677 GTCAGATTCTCAGCTTATAGAGCTGCCATGAACTCCGAAGACTGCAGAGGCCCTTTGC 3736
| | | | |
QY 1981 TTGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCCCTTGGACCAACAACCTCAAGCAA 2040
| | | | |
Db 3737 TTGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCCCTTGGACCAACAACCTCAAGCAA 3796
| | | | |
QY 2041 AATGACCAAGCCATGATATCTGACAGATTATTAATGTTGACCACTATTTATGACCGC 2100
| | | | |
Db 3797 AATGACCAAGCCATGATATCTGACAGATTATTAATGTTGACCACTATTTATGACCGC 3856
| | | | |
QY 2101 CTGAGCAAGAGACACAACAATTGGTCAACGCTCCCTCTGCGTGATATGTCTGAAC 2160
| | | | |
Db 3857 CTGAGCAAGAGACACAACAATTGGTCAACGCTCCCTCTGCGTGATATGTCTGAAC 3916
| | | | |
QY 2161 TGGCTGCTGAATGTTTATGATACGGGACGACAGAGGATCCGTGCTGCTTTTAA 2220
| | | | |
Db 3917 TGGCTGCTGAATGTTTATGATACGGGACGACAGAGGATCCGTGCTGCTTTTAA 3976
| | | | |
QY 2221 ACTGGCATATTTCCCTGTGTAAAGCACATTGGAGAACAAGTACAGATACCTTTTCAAG 2280
| | | | |
Db 3977 ACTGGCATATTTCCCTGTGTAAAGCACATTGGAGAACAAGTACAGATACCTTTTCAAG 4036
| | | | |
QY 2281 CAAGTGGCAAGTTCACACAGATTTTGTGACCAAGCGCAGGCTGGCCCTCTTCTGCATGAT 2340
| | | | |
Db 4037 CAAGTGGCAAGTTCACACAGATTTTGTGACCAAGCGCAGGCTGGCCCTCTTCTGCATGAT 4096
| | | | |
QY 2341 TCTATCCAAATTCCAAGACAGTTGGGTGAAGTGCATCTTTGGGGGCAAGTACATTGAG 2400
| | | | |
Db 4097 TCTATCCAAATTCCAAGACAGTTGGGTGAAGTGCATCTTTGGGGGCAAGTACATTGAG 4156
| | | | |
QY 2401 C 2401
|
Db 4157 C 4157

RESULT 3
US-09-845-416-35
; Sequence 35, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 4848
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-35
Query Match 100.0%; Score 2401; DB 13; Length 4848;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 2401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCAGTTCATTGATGAGAGAGTGAAGTAAACCTGGACCGGTATCAACAAGCTTTAGAGAA 60
| | | | |
Db 1780 GGCAGTTCATTGATGAGAGAGTGAAGTAAACCTGGACCGGTATCAACAAGCTTTAGAGAA 1839
| | | | |
QY 61 GATTTATCGTGGCTTCTTCTGCTGAGGACACTTGCAGACACAAGAGAGATTCTAAT 120
| | | | |
Db 1840 GATTTATCGTGGCTTCTTCTGCTGAGGACACTTGCAGACACAAGAGAGATTCTAAT 1899
| | | | |
QY 121 GATGTGAAGTGTGAAGAACCAAGTTTCATCTCATGAGGGGTACATGATGATTGACA 180
| | | | |
Db 1900 GATGTGAAGTGTGAAGAACCAAGTTTCATCTCATGAGGGGTACATGATGATTGACA 1959
| | | | |
QY 181 GCCCATCAGGGCCGGTGTGTAATATTCTCAATTTGGGAAGTGAAGCTGAATTGAAACAGA 240
| | | | |
Db 1960 GCCCATCAGGGCCGGTGTGTAATATTCTCAATTTGGGAAGTGAAGCTGAATTGAAACAGA 2019
| | | | |
QY 241 AATTTATCAGAGATGAAGAACTGAAGTACAGAGACAGATGAATCTCTAAATTCAAGA 300
| | | | |
Db 2020 AATTTATCAGAGATGAAGAACTGAAGTACAGAGACAGATGAATCTCTCTAAATTCAAGA 2079
| | | | |
QY 301 TGGGAATGCTCAGGGTACGTAGCATGGAAGAAACAAAGCAATTACATAGATTCTTAATG 360
| | | | |
Db 2080 TGGGAATGCTCAGGGTACGTAGCATGGAAGAAACAAAGCAATTACATAGATTCTTAATG 2139
| | | | |
QY 361 GATCTCCAGATCAGAAACTGAAGAGTGAATGACTGGCTAACAAAAACAAGAAAGA 420
| | | | |
Db 2140 GATCTCCAGATCAGAAACTGAAGAGTGAATGACTGGCTAACAAAAACAAGAAAGA 2199
| | | | |
QY 421 ACAAGAAATGAGAGAGAGCCCTCTTGACCTGATCTTGAAGACCTAAACGCCAAGTA 480
| | | | |
Db 2200 ACAAGAAATGAGAGAGAGCCCTCTTGACCTGATCTTGAAGACCTAAACGCCAAGTA 2259
| | | | |
QY 481 CAACAACATAAGTGTCTTCAAGAAAGATCTAGAACAAACAAGTCAAGGTCAATTCTCTC 540
| | | | |
Db 2260 CAACAACATAAGTGTCTTCAAGAAAGATCTAGAACAAACAAGTCAAGGTCAATTCTCTC 2319
| | | | |
QY 541 ACTCAATGCTGTGTGATGATGAATCTAGTGGAGATCAAGCAACTGCTCTTTGGA 600
| | | | |
Db 2320 ACTCAATGCTGTGTGATGATGAATCTAGTGGAGATCAAGCAACTGCTCTTTGGA 2379
| | | | |
QY 601 GAACAATTAAAGTATTGGAGATCGATGGGCAACATCTGTAGATGACAGAACACCGC 660
| | | | |
Db 2380 GAACAATTAAAGTATTGGAGATCGATGGGCAACATCTGTAGATGACAGAACACCGC 2439
| | | | |
QY 661 TGGGTCTTTTACAGACACTATAGATTACTGCAACAGTTCCCTGACCTGGAAG 720
| | | | |
Db 2440 TGGGTCTTTTACAGACACTATAGATTACTGCAACAGTTCCCTGACCTGGAAG 2499
| | | | |
QY 721 TTTCTGCTGGCTTACAGAGCTGAACCAACTGCAATGTCTACAGATGTAACCGT 780
| | | | |
Db 2500 TTTCTGCTGGCTTACAGAGCTGAACCAACTGCAATGTCTACAGATGTAACCGT 2559
| | | | |
QY 781 AAGGAAGCTCTAGAGACTCCAGGAGTAAAGAGCTGATGAACAACAATGGCAAGAC 840
| | | | |
Db 2560 AAGGAAGCTCTAGAGACTCCAGGAGTAAAGAGCTGATGAACAACAATGGCAAGAC 2619
| | | | |
QY 841 CTCGAAGTGAATTTGAAGCTCACACAGATGTTTATCAACAACCTGATGAACCAAGCCAA 900
| | | | |
Db 2620 CTCGAAGTGAATTTGAAGCTCACACAGATGTTTATCAACAACCTGATGAACCAAGCCAA 2679
| | | | |
QY 901 AAAATCTGAGATCCCTGGAAGTCCGATGATGCAAGTCTGTTCACAAAGACGTTGGAT 960
| | | | |
Db 2680 AAAATCTGAGATCCCTGGAAGTCCGATGATGCAAGTCTGTTCACAAAGACGTTGGAT 2739
| | | | |
QY 961 AACATGAATCTCAAGTGAAGTGAATTCGGAAAAAGTCTTCAACATTAGTCCCAATTG 1020
| | | | |
Db 2740 AACATGAATCTCAAGTGAAGTGAATTCGGAAAAAGTCTTCAACATTAGTCCCAATTG 2799
| | | | |
QY 1021 GAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACTTTCTCTGCAAGAACTCTGTGTGG 1080
| | | | |
Db 2800 GAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACTTTCTCTGCAAGAACTCTGTGTGG 2859
| | | | |


```
QY 1081 CTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCCCTATTGGAGCGACTTTCCAGCA 1140
Db 2860 CTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCCCTATTGGAGCGACTTTCCAGCA 2919
QY 1141 GTTCAGAACGAGACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAATAAGAACTT 1200
Db 2920 GTTCAGAACGAGACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAATAAGAACTT 2979
QY 1201 GTAATCATGAGTACTCTTGAGACTGTGACGAATATTTCTGACAGAGCAGCCCTTTGAAAGGA 1260
Db 2980 GTAATCATGAGTACTCTTGAGACTGTGACGAATATTTCTGACAGAGCAGCCCTTTGAAAGGA 3039
QY 1261 CTAGAGAAACTCTACAGAGAGCCCAAGAGAGCTGCCCTCTTGAGAGAGAGAGCCCAAGATGTC 1320
Db 3040 CTAGAGAAACTCTACAGAGAGCCCAAGAGAGCTGCCCTCTTGAGAGAGAGAGCCCAAGATGTC 3099
QY 1321 ACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCATATCTGAGTGGGAAAAATTGAACCTG 1380
Db 3100 ACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCATATCTGAGTGGGAAAAATTGAACCTG 3159
QY 1381 CACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAA 1440
Db 3160 CACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAA 3219
QY 1441 GAGGCCACGATGAGCTGGAACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGAATCCTGG 1500
Db 3220 GAGGCCACGATGAGCTGGAACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGAATCCTGG 3279
QY 1501 CAGCCCGTGGGCGATCTCCTCATGTACTCTCTCCAAGATCACTCGAGAAAGTCAAGGCA 1560
Db 3280 CAGCCCGTGGGCGATCTCCTCATGTACTCTCTCCAAGATCACTCGAGAAAGTCAAGGCA 3339
QY 1561 CTTGAGAGAGAAATTGCGCCTCTGAAAGAGAGACGTGAGCCACGTCAATGACCTTGCTCGC 1620
Db 3340 CTTGAGAGAGAAATTGCGCCTCTGAAAGAGAGACGTGAGCCACGTCAATGACCTTGCTCGC 3399
QY 1621 CAGCTTACCACTTTGGGCATTCAGTCTCACCGTATACTCAGCACTCTGGAAGACCTG 1680
Db 3400 CAGCTTACCACTTTGGGCATTCAGTCTCACCGTATACTCAGCACTCTGGAAGACCTG 3459
QY 1681 AACACCAAGATGGAAGCTTCTGCAAGTGGCCGCTGAGAGACCGAGTCAGGCAAGTGAAGAA 1740
Db 3460 AACACCAAGATGGAAGCTTCTGCAAGTGGCCGCTGAGAGACCGAGTCAGGCAAGTGAAGAA 3519
QY 1741 GCCCAACAGGAACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTGTGCCAGGGTCCC 1800
Db 3520 GCCCAACAGGAACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTGTGCCAGGGTCCC 3579
QY 1801 TGGGAGAGAGCCATCTGCGCAAAAGAGTGCCTTACTATATCAACCAAGAGACTCAAAACA 1860
Db 3580 TGGGAGAGAGCCATCTGCGCAAAAGAGTGCCTTACTATATCAACCAAGAGACTCAAAACA 3639
QY 1861 ACTTGCTGGACCATCCCAAAATGACAGAGCTTACAGTCTTTAGCTGACCTGAATAAT 1920
Db 3640 ACTTGCTGGACCATCCCAAAATGACAGAGCTTACAGTCTTTAGCTGACCTGAATAAT 3699
QY 1921 GTCAGATTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAGGCGCCCTTTGC 1980
Db 3700 GTCAGATTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAGGCGCCCTTTGC 3759
QY 1981 TTGATCTCTGAGCCTGTGAGCTGATGTGATGCTTGACCAAGCAACCTCAAGCAA 2040
Db 3760 TTGATCTCTGAGCCTGTGAGCTGATGTGATGCTTGACCAAGCAACCTCAAGCAA 3819
QY 2041 AATGACCAAGCCATGATATCTGAGATTATTAATTGTTGACCACTATTATGACCGC 2100
Db 3820 AATGACCAAGCCATGATATCTGAGATTATTAATTGTTGACCACTATTATGACCGC 3879
QY 2101 CTGAGAGAGAGCAACAATTTGTTCAAGCTCCCTCTCTGGGTGATATGTGTCTGAAC 2160
Db 3880 CTGAGAGAGAGCAACAATTTGTTCAAGCTCCCTCTCTGGGTGATATGTGTCTGAAC 3939
```

```
QY 2161 TGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGATCCGTCTGTCTTTAAA 2220
Db 3940 TGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGATCCGTCTGTCTTTAAA 3999
QY 2221 ACTGGCATATTTCCCTGTGTAAAGCAATTGGAAGACAGTACAGATACCTTTTCAAG 2280
Db 4000 ACTGGCATATTTCCCTGTGTAAAGCAATTGGAAGACAGTACAGATACCTTTTCAAG 4059
QY 2281 CAAGTGGCAAGTTCAACAGAGATTTGTGACCAAGCGCAGGCTGGGCTCTCTGCAATGAT 2340
Db 4060 CAAGTGGCAAGTTCAACAGAGATTTGTGACCAAGCGCAGGCTGGGCTCTCTGCAATGAT 4119
QY 2341 TCTATCCAATTTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGCAGTAACATTTAG 2400
Db 4120 TCTATCCAATTTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGCAGTAACATTTAG 4179
QY 2401 C 2401
Db 4180 C 4180
```

```
RESULT 4
US-09-845-416-36
; Sequence 36, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 5060
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-36
```

```
Query Match 100.0%; Score 2401; DB 13; Length 5060;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCAGTTCATTTGATGAGAGTGAAGTAAACCTGGAACCGTTATCAAAACAGCTTTAGAGAA 60
Db 1992 GGCAGTTCATTTGATGAGAGTGAAGTAAACCTGGAACCGTTATCAAAACAGCTTTAGAGAA 2051
QY 61 GATTTATCGTGGCTTCTTCTGCTGAGACACATTTGCAAGCAACAAGAGAGATTTCTAAT 120
Db 2052 GATTTATCGTGGCTTCTTCTGCTGAGACACATTTGCAAGCAACAAGAGAGATTTCTAAT 2111
QY 121 GATGTGAAGTGTGAAGAAGCAAGTTTCATATCTCATGAGGGGTACATGATGATTTGACA 180
Db 2112 GATGTGAAGTGTGAAGAAGCAAGTTTCATATCTCATGAGGGGTACATGATGATTTGACA 2171
QY 181 GCCCATCAGGGCCGGGTTGTAATATTTCTACAATTTGGGAAGTAAGCTGATTGGAACAGA 240
Db 2172 GCCCATCAGGGCCGGGTTGTAATATTTCTACAATTTGGGAAGTAAGCTGATTGGAACAGA 2231
QY 241 AATTATCAGAAGATGAAGAAGTGAAGTACAGAGACAGATGAATCTCCTAATTTCAAGA 300
Db 2232 AATTATCAGAAGATGAAGAAGTGAAGTACAGAGACAGATGAATCTCCTAATTTCAAGA 2291
QY 301 TGGGAATGCTCAGGGTAGCTAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATG 360
Db 2292 TGGGAATGCTCAGGGTAGCTAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATG 2351
QY 361 GATCTCCAGAAATCAGAACTGAAGAAGTGAATGACTGGCTTAACAAAAACAGAGAAAGA 420
Db 2352 GATCTCCAGAAATCAGAACTGAAGAAGTGAATGACTGGCTTAACAAAAACAGAGAAAGA 2411
```


| | | | |
|----|------|---|------|
| QY | 421 | ACAAAGAAATGGAAGAGAAGCCTCTTGACCTGATCTTTGAAGACCTTAAACGCCAAGTA | 480 |
| Db | 2412 | ACAAAGAAATGGAAGAGAAGCCTCTTGACCTGATCTTTGAAGACCTTAAACGCCAAGTA | 2471 |
| QY | 481 | CAACAACATTAAGTGCTTCAAGAAGATCTAGAACAGAACAAAGTCAGGGGTCAATTCTCTC | 540 |
| Db | 2472 | CAACAACATTAAGTGCTTCAAGAAGATCTAGAACAGAACAAAGTCAGGGGTCAATTCTCTC | 2531 |
| QY | 541 | ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAACGCACTGCTGCTTTGGAA | 600 |
| Db | 2532 | ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAACGCACTGCTGCTTTGGAA | 2591 |
| QY | 601 | GAACAACCTTAAGGTAATGGGAGATCGATGGGCAACAATCTGTAGATGGAACAGAACCCG | 660 |
| Db | 2592 | GAACAACCTTAAGGTAATGGGAGATCGATGGGCAACAATCTGTAGATGGAACAGAACCCG | 2651 |
| QY | 661 | TGGGTTCTTTTACAAGACATCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAAG | 720 |
| Db | 2652 | TGGGTTCTTTTACAAGACATCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAAG | 2711 |
| QY | 721 | TTTCTTGCCCTGGCTTACAGAAGCTGAAACAACCTGCCAATGTCTTACAGAGATGCTACCCGT | 780 |
| Db | 2712 | TTTCTTGCCCTGGCTTACAGAAGCTGAAACAACCTGCCAATGTCTTACAGAGATGCTACCCGT | 2771 |
| QY | 781 | AAGGAAAGGCTCCTAGAGACTCCAAGGGAGTAAAGAAGCTGATGAAACAATGGCAAGAC | 840 |
| Db | 2772 | AAGGAAAGGCTCCTAGAGAAGCTCCAAGGGAGTAAAGAAGCTGATGAAACAATGGCAAGAC | 2831 |
| QY | 841 | CTCCAAGGTGAATGGAAGCTCACACAGATGTTTATCACAACCTGATGAAAAACGCCAA | 900 |
| Db | 2832 | CTCCAAGGTGAATGGAAGCTCACACAGATGTTTATCACAACCTGATGAAAAACGCCAA | 2891 |
| QY | 901 | AAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAAGTCTGTTCACAAGAAGTTTGAT | 960 |
| Db | 2892 | AAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAAGTCTGTTCACAAGAAGTTTGAT | 2951 |
| QY | 961 | AACATGAACCTTCAAGTGAAGTGAAGTTCGAAAAAGTCTCTCAACATTAGTCCCAATTG | 1020 |
| Db | 2952 | AACATGAACCTTCAAGTGAAGTGAAGTTCGAAAAAGTCTCTCAACATTAGTCCCAATTG | 3011 |
| QY | 1021 | GAAAGCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGACAGAACTTCTGGTGG | 1080 |
| Db | 3012 | GAAAGCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGACAGAACTTCTGGTGG | 3071 |
| QY | 1081 | CTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACTATTGGAGGCACTTCCAGCA | 1140 |
| Db | 3072 | CTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACTATTGGAGGCACTTCCAGCA | 3131 |
| QY | 1141 | GTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAGAAGCT | 1200 |
| Db | 3132 | GTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAGAAGCT | 3191 |
| QY | 1201 | GTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGACAGCCTTTGGAAGA | 1260 |
| Db | 3192 | GTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGACAGCCTTTGGAAGA | 3251 |
| QY | 1261 | CTAGAGAAACTCTACCAGAGGCCAGAGAGCTGCCTCCTGAGGAGAGAGGCCAGAAATGTC | 1320 |
| Db | 3252 | CTAGAGAAACTCTACCAGAGGCCAGAGAGCTGCCTCCTGAGGAGAGAGGCCAGAAATGTC | 3311 |
| QY | 1321 | ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATTACTGAGTGGGAAAAATTGAACCTG | 1380 |
| Db | 3312 | ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATTACTGAGTGGGAAAAATTGAACCTG | 3371 |
| QY | 1381 | CACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGAAGCTCCAAGAACTTCAA | 1440 |
| Db | 3372 | CACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGAAGCTCCAAGAACTTCAA | 3431 |
| QY | 1441 | GAGGCCACGGATGAGCTGGAAGCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGATCTCTGG | 1500 |
| Db | 3432 | GAGGCCACGGATGAGCTGGAAGCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGATCTCTGG | 3491 |

| | | | |
|----|------|---|------|
| QY | 1501 | CAGCCCCGTGGGGGATCTCCCTCATTTGACTCTCTCCAAAGATCAACCTCGAGAAAGTCAAGGCA | 1560 |
| Db | 3492 | CAGCCCCGTGGGGGATCTCCCTCATTTGACTCTCTCCAAAGATCAACCTCGAGAAAGTCAAGGCA | 3551 |
| QY | 1561 | CTTCGAGGAGAAAAATTGGCCCTCTGAAGAAGAAAGTGAGCCACGTCATGACTCTTGCTCGC | 1620 |
| Db | 3552 | CTTCGAGGAGAAAAATTGGCCCTCTGAAGAAGAAAGTGAGCCACGTCATGACTCTTGCTCGC | 3611 |
| QY | 1621 | CAGCTTACCATTCTGGGGCATTCAGCTCTCAACCGTATTAACCTCAGCACTCTGGAGAAGACCTG | 1680 |
| Db | 3612 | CAGCTTACCATTCTGGGGCATTCAGCTCTCAACCGTATTAACCTCAGCACTCTGGAGAAGACCTG | 3671 |
| QY | 1681 | AACACCCAGATGGAAGCTTCTGCAGAGTGGCCGTGAGGAACCGAGTCAGGCAGCTGCATGAA | 1740 |
| Db | 3672 | AACACCCAGATGGAAGCTTCTGCAGAGTGGCCGTGAGGAACCGAGTCAGGCAGCTGCATGAA | 3731 |
| QY | 1741 | GCCCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCACAGCTGTGTCAGGGTCCC | 1800 |
| Db | 3732 | GCCCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCACAGCTGTGTCAGGGTCCC | 3791 |
| QY | 1801 | TGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACACGAGACTCAAAACA | 1860 |
| Db | 3792 | TGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACACGAGACTCAAAACA | 3851 |
| QY | 1861 | ACTTGCTGGAGCCATCCCAAATGACAGAGCTCTACCACTTTAGCTGACTGAATTAAT | 1920 |
| Db | 3852 | ACTTGCTGGAGCCATCCCAAATGACAGAGCTCTACCACTTTAGCTGACTGAATTAAT | 3911 |
| QY | 1921 | GTCAGATTTCTAGCTTATAGGACTGCCATGAAATCCGAAGACTGCAGAAAGCCCTTTC | 1980 |
| Db | 3912 | GTCAGATTTCTAGCTTATAGGACTGCCATGAAATCCGAAGACTGCAGAAAGCCCTTTC | 3971 |
| QY | 1981 | TTGGATCTCTTGAGCCTGTCAAGTGCATGTGATGCTTGAGCCAGACAACCTCAAGCAA | 2040 |
| Db | 3972 | TTGGATCTCTTGAGCCTGTCAAGTGCATGTGATGCTTGAGCCAGACAACCTCAAGCAA | 4031 |
| QY | 2041 | AATGACCAGCCCATGGATATCTCGAGATTAATTAATGTTTGACCACATAATTAAGACCGC | 2100 |
| Db | 4032 | AATGACCAGCCCATGGATATCTCGAGATTAATTAATGTTTGACCACATAATTAAGACCGC | 4091 |
| QY | 2101 | CTGAGCAAGAGACAACAATTTGTCAACGTCCCTCTCGCTGATATGTCTGAAC | 2160 |
| Db | 4092 | CTGAGCAAGAGACAACAATTTGTCAACGTCCCTCTCGCTGATATGTCTGAAC | 4151 |
| QY | 2161 | TGGCTGCTGAATTTTATGATACGGAGCAAGAGGAGATCCGTCTCTGCTTTTAA | 2220 |
| Db | 4152 | TGGCTGCTGAATTTTATGATACGGAGCAAGAGGAGATCCGTCTCTGCTTTTAA | 4211 |
| QY | 2221 | ACTGGCATTTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAG | 2280 |
| Db | 4212 | ACTGGCATTTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAG | 4271 |
| QY | 2281 | CAAGTGGAAGTCAACAGGATTTTGTGACCAAGCCAGGCTGGGCTCCTTCTGCATGAT | 2340 |
| Db | 4272 | CAAGTGGAAGTCAACAGGATTTTGTGACCAAGCCAGGCTGGGCTCCTTCTGCATGAT | 4331 |
| QY | 2341 | TCTATCCAAATTCACAACAAGTGGGTGAAGTGCATCCTTTGGGGGCAAGTAACATTGAG | 2400 |
| Db | 4332 | TCTATCCAAATTCACAACAAGTGGGTGAAGTGCATCCTTTGGGGGCAAGTAACATTGAG | 4391 |
| QY | 2401 | C 2401 | |
| Db | 4392 | C 4392 | |

RESULT 5
 US-09-845-416-6
 ; Sequence 6, Application US/09845416
 ; Publication No. US20030171312A1
 ; GENERAL INFORMATION:
 ; APPLICANT: XIAO, XIAO
 ; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
 ; TITLE OF INVENTION: THEREOF

FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 3999
TYPE: DNA
ORGANISM: Homo sapiens
US-09-845-416-6

Query Match 93.7%; Score 2250; DB 13; Length 3999;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 141; Gaps 1;

QY 1 GGCAGTTCATTGATGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAA 60
DB 1000 GGCAGTTCATTGATGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAA 1059
QY 61 GTATTATCGTGCTCTTTCTGCTGAGACACATTGCAAGCACAAGAGAGATTCTAAT 120
DB 1060 GTATTATCGTGCTCTTTCTGCTGAGACACATTGCAAGCACAAGAGAGATTCTAAT 1119
QY 121 GATGTGAAGTGTGTAAGAACCACTTCTATCTCATGAGGGTACATGATGATTGACA 180
DB 1120 GATGTGAAGTGTGTAAGAACCACTTCTATCTCATGAGGGTACATGATGATTGACA 1179
QY 181 GCCCATCAGGGCCGGGTTGTAATATTCTACAATTGGAGTAAGCTGATTGGACAGGA 240
DB 1180 GCCCATCAGGGCCGGGTTGTAATATTCTACAATTGGAGTAAGCTGATTGGACAGGA 1239
QY 241 AAATTATCAGAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTGAGA 300
DB 1240 AAATTATCAGAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTGAGA 1299
QY 301 TGGGAATGCTCAGGGTAGCTAGCTAGGAAAAAACAAGCAATTTACATAGATTTAATG 360
DB 1300 TGGGAATGCTCAGGGTAGCTAGCTAGGAAAAAACAAGCAATTTACATAGATTTAATG 1359
QY 361 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAGAAAGA 420
DB 1360 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAGAAAGA 1419
QY 421 ACAAGGAAATGAGAGAGAGCTCTTGGAACCTGATCTTGAGAGCCTAAAAACGCCAAGTA 480
DB 1420 ACAAGGAAATGAGAGAGAGCTCTTGGAACCTGATCTTGAGAGCCTAAAAACGCCAAGTA 1479
QY 481 CAACAACATAAGGTGCTTCAAGAGATCTAGAACAAAGAACAGTCAAGGCTCAATTCTCTC 540
DB 1480 CAACAACATAAGGTGCTTCAAGAGATCTAGAACAAAGAACAGTCAAGGCTCAATTCTCTC 1539
QY 541 ACTCACATGCTGTGTAGTGTGATGAATCTAGTGAGATCAAGCACTGCTGTTTGAA 600
DB 1540 ACTCACATGCTGTGTAGTGTGATGAATCTAGTGAGATCAAGCACTGCTGTTTGAA 1599
QY 601 GAACAACCTTAAGTATTGGAGATCGATGGCAAAACATCTGTAGTGACAGAAACCGC 660
DB 1600 GAACAACCTTAAGTATTGGAGATCGATGGCAAAACATCTGTAGTGACAGAAACCGC 1659
QY 661 TGGGTTCTTTTACAAGAC----- 678
DB 1660 TGGGTTCTTTTACAAGACCGCTGACCTAGCTCCTGGACTGACCACTATTGGACCTCT 1719
QY 679 ----- 678
DB 1720 CCTACTCAGACTGTACTCTGTGTACACAACTGTGTACTTAAGGAACCTGCCATCTCC 1779
QY 679 -----ACTCATAGATTACTGCAACAG 699
DB 1780 AAACCTAGAAATGCCATCTTCTTGTGATTGAGGTACTACTCATAGATTACTGCAACAG 1839

QY 700 TTCCCCCTGACCTTGAAAAAGTTTCTGCTGCTTACAGAGCTGAAACAACCTGCCAAT 759
DB 1840 TTCCCCCTGACCTTGAAAAAGTTTCTGCTGCTTACAGAGCTGAAACAACCTGCCAAT 1899
QY 760 GTCCCTACAGATGCTACCCGTAAGGAAAGCTCCTAGAGAACTCCAAGGAGTAAGAG 819
DB 1900 GTCCCTACAGATGCTACCCGTAAGGAAAGCTCCTAGAGAACTCCAAGGAGTAAGAG 1959
QY 820 CTGATGAACAATGGCAAGACCTCCAGAGTGAATTTGAAGCTCACACAGATGTTATCAC 879
DB 1960 CTGATGAACAATGGCAAGACCTCCAGAGTGAATTTGAAGCTCACACAGATGTTATCAC 2019
QY 880 AACCTGATGAACCAAGCCCAAAAAATCCTGAGATCCCTGGAAGTTCCGATGATGACATC 939
DB 2020 AACCTGATGAACCAAGCCCAAAAAATCCTGAGATCCCTGGAAGTTCCGATGATGACATC 2079
QY 940 CTGTTACAAGACGTTTGATTAACATGAACCTTCAAGTGAAGTGAACCTTCGAAAAAGTCT 999
DB 2080 CTGTTACAAGACGTTTGATTAACATGAACCTTCAAGTGAAGTGAACCTTCGAAAAAGTCT 2139
QY 1000 CTCAACATTAAGTCCCATTTGGAAGCCAGTTCTGACCAGTGAAGCGCTGCACTTTCT 1059
DB 2140 CTCAACATTAAGTCCCATTTGGAAGCCAGTTCTGACCAGTGAAGCGCTGCACTTTCT 2199
QY 1060 CTGCAGAACTTCTGTGTGCTACAGCTGAAGAGATGATGAATTAAGCCGGCAGGACCT 1119
DB 2200 CTGCAGAACTTCTGTGTGCTACAGCTGAAGAGATGATGAATTAAGCCGGCAGGACCT 2259
QY 1120 ATTGAGGCACTTTCAGACAGTTCAGAAACAGACGATGATGATGAGGCTTCAAGAG 1179
DB 2260 ATTGAGGCACTTTCAGACAGTTCAGAAACAGACGATGATGATGAGGCTTCAAGAG 2319
QY 1180 GAATTGAACCTAAGAACTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTG 1239
DB 2320 GAATTGAACCTAAGAACTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTG 2379
QY 1240 ACAGAGACCTTTGGAAGACTAGAGAACTCTACAGAGCCCAAGAGCTGCTCCT 1299
DB 2380 ACAGAGACCTTTGGAAGACTAGAGAACTCTACAGAGCCCAAGAGCTGCTCCT 2439
QY 1300 GAGAGAGACCCCAAGATGTCACTCCGCTTCTACGAAGCAGCTGAGAGGTCAATACT 1359
DB 2440 GAGAGAGACCCCAAGATGTCACTCCGCTTCTACGAAGCAGCTGAGAGGTCAATACT 2499
QY 1360 GAGTGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTT 1419
DB 2500 GAGTGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTT 2559
QY 1420 GAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGCAAGCTGCCCAAGCT 1479
DB 2560 GAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGCAAGCTGCCCAAGCT 2619
QY 1480 GAGGTGATCAAGGATCTCTGGAGCCCGTGGCGATCTCTCATTTGACTCTCTCCAGAT 1539
DB 2620 GAGGTGATCAAGGATCTCTGGAGCCCGTGGCGATCTCTCATTTGACTCTCTCCAGAT 2679
QY 1540 CACCTGAGAGAAAGTCAAGGCACTTCAAGAGAGAAATTGCGCTCTGAAAGAGAACGTGAGC 1599
DB 2680 CACCTGAGAGAAAGTCAAGGCACTTCAAGAGAGAAATTGCGCTCTGAAAGAGAACGTGAGC 2739
QY 1600 CACGTCATGACCTTGTGCGCAGCTTACCACTTTGGGCACTCAGCTCTCACCGTATTAAC 1659
DB 2740 CACGTCATGACCTTGTGCGCAGCTTACCACTTTGGGCACTCAGCTCTCACCGTATTAAC 2799
QY 1660 CTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCAAGTGGCCGTCAGAGAC 1719
DB 2800 CTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCAAGTGGCCGTCAGAGAC 2859
QY 1720 CGAGTCAGGCACTGTCATGAGCCCAAGGCACTTGGTCCAGCATCTCAGACCTTTCTT 1779
DB 2860 CGAGTCAGGCACTGTCATGAGCCCAAGGCACTTGGTCCAGCATCTCAGACCTTTCTT 2919
QY 1780 TCCAGCTCTGTCCAGGCTCCCTGGAGAGAGACCATCTGCCAAACAAAGTGCCCTACTAT 1839

Db 2920 TCCACGCTCTGCCAGGCTCCCTGGAGAGAGCCATCTCGCCAAACAAGTGCCCTACTAT 2979
QY 1840 ATCAACCAAGAGACTCAAAACAACCTTGCTGGAGCATCCCAAAATGACAGAGCTTACCAG 1899
Db 2980 ATCAACCAAGAGACTCAAAACAACCTTGCTGGAGCATCCCAAAATGACAGAGCTTACCAG 3039
QY 1900 TCTTAGCTGACCTGAAATATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGA 1959
Db 3040 TCTTAGCTGACCTGAAATATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGA 3099
QY 1960 AGACTGAGAGAGGCCCTTTGCTTGATCTCTTGAGCCTGTGAGCTGCATGTGATGCTTG 2019
Db 3100 AGACTGAGAGAGGCCCTTTGCTTGATCTCTTGAGCCTGTGAGCTGCATGTGATGCTTG 3159
QY 2020 GACCAGCAACCTCAAGCAAAATGACAGCCCATGATATCTCGAGATTATTAATTGT 2079
Db 3160 GACCAGCAACCTCAAGCAAAATGACAGCCCATGATATCTCGAGATTATTAATTGT 3219
QY 2080 TTGACCACTATTTATGACCGCCTGAGCAAGACACAATTTGGTCAACGCTCCTCTC 2139
Db 3220 TTGACCACTATTTATGACCGCCTGAGCAAGACACAATTTGGTCAACGCTCCTCTC 3279
QY 2140 TGGGTGATATGTGTGTAAGTGTGTAAGTGTGTAAGTGTGTAAGTGTGTAAGTGTGTAAG 2199
Db 3280 TGGGTGATATGTGTGTAAGTGTGTAAGTGTGTAAGTGTGTAAGTGTGTAAGTGTGTAAG 3339
QY 2200 ATCCGTGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGAC 2259
Db 3340 ATCCGTGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGAC 3399
QY 2260 AAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGAGATTTTGACCAAGCCAGG 2319
Db 3400 AAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGAGATTTTGACCAAGCCAGG 3459
QY 2320 CTGGGCTTCTTCTGCATGATTTCTCAAAATTTCCAGAGAGTGGGTGAAGTTGATCC 2379
Db 3460 CTGGGCTTCTTCTGCATGATTTCTCAAAATTTCCAGAGAGTGGGTGAAGTTGATCC 3519
QY 2380 TTTGGGGGAGTAACATTGAGC 2401
Db 3520 TTTGGGGGAGTAACATTGAGC 3541

RESULT 6

US-09-845-416-28
; Sequence 28, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 4966
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-28

Query Match 93.7%; Score 2250; DB 13; Length 4966;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 141; Gaps 1;
QY 1 GGCAGTTCATTCATGAGAGAGTGAAGTAACCTTGACCGTTATCAACAGCTTTAGAGAA 60
Db 1757 GGCAGTTCATTCATGAGAGAGTGAAGTAACCTTGACCGTTATCAACAGCTTTAGAGAA 1816

QY 61 GTATTATCGTGCTTCTTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTCTAAT 120
Db 1817 GTATTATCGTGCTTCTTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTCTAAT 1876
QY 121 GATGTGAAGTGTGAAAGACAGTTTCATCTCATGAGGGGTACATGATGATTGACA 180
Db 1877 GATGTGAAGTGTGAAAGACAGTTTCATCTCATGAGGGGTACATGATGATTGACA 1936
QY 181 GCCCATCAGGGCCGGGTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGAAACAGGA 240
Db 1937 GCCCATCAGGGCCGGGTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGAAACAGGA 1996
QY 241 AATTATCAGAAGATGAAGAACTGAAGTACAAAGAGAGATGAATCTCCTAAATTCAAGA 300
Db 1997 AATTATCAGAAGATGAAGAACTGAAGTACAAAGAGAGATGAATCTCCTAAATTCAAGA 2056
QY 301 TGGGAATGCTCAGGGTAGCTAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATG 360
Db 2057 TGGGAATGCTCAGGGTAGCTAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATG 2116
QY 361 GATCTCCAGATCAGAAACTGAAGAAGTTGAATGACTGGCTAACAAACAGAGAAAGA 420
Db 2117 GATCTCCAGATCAGAAACTGAAGAAGTTGAATGACTGGCTAACAAACAGAGAAAGA 2176
QY 421 ACAAGGAAATGAGAGAGAGCTCTTGACCTGATCTTGAAGACCTAAACGCCAAGTA 480
Db 2177 ACAAGGAAATGAGAGAGAGCTCTTGACCTGATCTTGAAGACCTAAACGCCAAGTA 2236
QY 481 CAACAACATTAAGTGTCTTCAAGAAAGATCTGAACAACAAGCAAGTCAAGGTCATTTCTC 540
Db 2237 CAACAACATTAAGTGTCTTCAAGAAAGATCTGAACAACAAGCAAGTCAAGGTCATTTCTC 2296
QY 541 ACTCACATGCTGTGTGATGATGTAATCTAGTGAGATCAAGCAACTGCTGCTTTGAA 600
Db 2297 ACTCACATGCTGTGTGATGATGTAATCTAGTGAGATCAAGCAACTGCTGCTTTGAA 2356
QY 601 GAACAACCTTAAGGTATTGGAGATGATGGCAACATCTGTAGATGACAGAGACCCG 660
Db 2357 GAACAACCTTAAGGTATTGGAGATGATGGCAACATCTGTAGATGACAGAGACCCG 2416
QY 661 TGGGTCTTTTACAAGAC----- 678
Db 2417 TGGGTCTTTTACAAGAC----- 2476
QY 679 ----- 678
Db 2477 CCTACTCAGACTGTACTCTGTGACACAACCTGTGTTACTAAGAACTGCCATCTCC 2536
QY 679 ----- 678
Db 2537 AAACTAGAATGCCATCTTCTTGATGTTGAGGTACTACTCATAGATTACTGACAG 2596
QY 700 TTCCCTGAGACCTGGAAGTTTCTGCTGCTGCTTACAGAACTGAAACAACCTGCCAAT 759
Db 2597 TTCCCTGAGACCTGGAAGTTTCTGCTGCTGCTTACAGAACTGAAACAACCTGCCAAT 2656
QY 760 GTCTACAGATGCTACCCGTAAGGAAGGCTCCTAGAGACTCCAAGGAGTAAGAG 819
Db 2657 GTCTACAGATGCTACCCGTAAGGAAGGCTCCTAGAGACTCCAAGGAGTAAGAG 2716
QY 820 CTGATGAACAATGGCAAGACCTTCAAGGTGAATTTGAAGCTCACACAGATGTTATCAC 879
Db 2717 CTGATGAACAATGGCAAGACCTTCAAGGTGAATTTGAAGCTCACACAGATGTTATCAC 2776
QY 880 AACCTGATGAAGAAACAGCCAAATTCCTGAGATCCCTGGAAGGTTCCGATGATGCAATC 939
Db 2777 AACCTGATGAAGAAACAGCCAAATTCCTGAGATCCCTGGAAGGTTCCGATGATGCAATC 2836
QY 940 CTGTTACAAAGACGTTTGATTAACATGAACCTTCAAGTGAAGTGAAGTTCGGAAGAAAGTCT 999
Db 2837 CTGTTACAAAGACGTTTGATTAACATGAACCTTCAAGTGAAGTGAAGTTCGGAAGAAAGTCT 2896
QY 1000 CTCAACATTAGGCTCCCATTTGGAAGCCAGTTCTGACCAAGTGAAGCGCTGCACTTTCT 1059

Db 2897 CTCAACATATAGGTCCCATTTGGAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTCT 2956
QY 1060 CTGCAGAACTTCTGTGTGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCT 1119
Db 2957 CTGCAGAACTTCTGTGTGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCT 3016
QY 1120 ATTGAGGCACTTTCAGCAGTTTCAGAGCAGAACGATGTACATAGGGCTTCAAGAG 1179
Db 3017 ATTGAGGCACTTTCAGCAGTTTCAGAGCAGAACGATGTACATAGGGCTTCAAGAG 3076
QY 1180 GAATTGAATACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 1239
Db 3077 GAATTGAATACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 3136
QY 1240 ACAGAGCAGCTTTGGAGAGCTAGAGAACTCTACAGAGCCCAAGAGAGCTGCTCCT 1299
Db 3137 ACAGAGCAGCTTTGGAGAGCTAGAGAACTCTACAGAGCCCAAGAGAGCTGCTCCT 3196
QY 1300 GAGGAGAGCCCAAGATGTCACTGCTTCTACGAAAGCAGGCTGAGAGGTCAATACT 1359
Db 3197 GAGGAGAGCCCAAGATGTCACTGCTTCTACGAAAGCAGGCTGAGAGGTCAATACT 3256
QY 1360 GAGTGGAAAAATGAACTGCACTCCGCTGACTGGCAGAAAAATAGATGACCCCTT 1419
Db 3257 GAGTGGAAAAATGAACTGCACTCCGCTGACTGGCAGAAAAATAGATGACCCCTT 3316
QY 1420 GAAAGACTCCAGGAACCTTCAAGAGCCACGATGAGCTGACCTCAAGCTGGCCAACT 1479
Db 3317 GAAAGACTCCAGGAACCTTCAAGAGCCACGATGAGCTGACCTCAAGCTGGCCAACT 3376
QY 1480 GAGGTGATCAAGGGATCCTGGCAGCCCTGGCGATCTCCTCATTTGACTCTTCCAGAT 1539
Db 3377 GAGGTGATCAAGGGATCCTGGCAGCCCTGGCGATCTCCTCATTTGACTCTTCCAGAT 3436
QY 1540 CACCTCGAAGAACTCAAGGCACTTCAGAGAGAAATTCGCTCTGAAGAGAACTGAGC 1599
Db 3437 CACCTCGAAGAACTCAAGGCACTTCAGAGAGAAATTCGCTCTGAAGAGAACTGAGC 3496
QY 1600 CACGTCAATGACCTTGTCTGCCAGCTTACCCTTTGGGCAATTCAGCTCTCACCCTAAT 1659
Db 3497 CACGTCAATGACCTTGTCTGCCAGCTTACCCTTTGGGCAATTCAGCTCTCACCCTAAT 3556
QY 1660 CTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAAGTGGCCCTGAGGAC 1719
Db 3557 CTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAAGTGGCCCTGAGGAC 3616
QY 1720 CGAGTACGGCAGCTGCATGAAGCCCAAGGACTTGTCCAGCATCTCAGCACTTTCTT 1779
Db 3617 CGAGTACGGCAGCTGCATGAAGCCCAAGGACTTGTCCAGCATCTCAGCACTTTCTT 3676
QY 1780 TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTGCCAAACAAAGTGCCTACTAT 1839
Db 3677 TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTGCCAAACAAAGTGCCTACTAT 3736
QY 1840 ATCAACGACGAGACTCAAACTTGTGCGGACCATCCCAAAATGACAGAGCTTACAG 1899
Db 3737 ATCAACGACGAGACTCAAACTTGTGCGGACCATCCCAAAATGACAGAGCTTACAG 3796
QY 1900 TCTTAGCTGACCTGAATAATGTGAGTTCTCAGCTTATAGGACTGCCATGAATCCGA 1959
Db 3797 TCTTAGCTGACCTGAATAATGTGAGTTCTCAGCTTATAGGACTGCCATGAATCCGA 3856
QY 1960 AGACTGAGAGAGCCCTTGTGCTGATCTCTGAGCCTGTCAGCTGCATGTGATGCTTG 2019
Db 3857 AGACTGAGAGAGCCCTTGTGCTGATCTCTGAGCCTGTCAGCTGCATGTGATGCTTG 3916
QY 2020 GACCAAGCAACCTCAAGCAAAATGACCAAGCCCATGATATCCTGCAGATTATTAATTGT 2079
Db 3917 GACCAAGCAACCTCAAGCAAAATGACCAAGCCCATGATATCCTGCAGATTATTAATTGT 3976
QY 2080 TTGACCACTATTATGACCGCTGAGCAAGAGCAACAATTGTCACGTCCTCTC 2139

Db 3977 TTGACCACTATTATGACCGCTGAGCAAGAGCAACAATTGCTCAACGTCCTCTC 4036
QY 2140 TCGGTGATATGTGTCTGAACCTGGCTGTAATGTTATGATACGGGACGAACAGGAG 2199
Db 4037 TCGGTGATATGTGTCTGAACCTGGCTGTAATGTTATGATACGGGACGAACAGGAG 4096
QY 2200 ATCCGTCTCTGCTTTTAAACCTGCATCATTTCCCTGTGTAAAGCAGCATTTGGAAGC 2259
Db 4097 ATCCGTCTCTGCTTTTAAACCTGCATCATTTCCCTGTGTAAAGCAGCATTTGGAAGC 4156
QY 2260 AAGTACAGATACCTTTTCAAGCAAGTGCAAGTTCAACAGGATTTTGTGACGAGCAGG 2319
Db 4157 AAGTACAGATACCTTTTCAAGCAAGTGCAAGTTCAACAGGATTTTGTGACGAGCAGG 4216
QY 2320 CTGGGCTCTCTTCTGCATGATTTCAATCCAAATTCCAAGAGAGTTGGGTGAAGTTGATCC 2379
Db 4217 CTGGGCTCTCTTCTGCATGATTTCAATCCAAATTCCAAGAGAGTTGGGTGAAGTTGATCC 4276
QY 2380 TTTGGGGCAGTAACATTGAGC 2401
Db 4277 TTTGGGGCAGTAACATTGAGC 4298

RESULT 7

US-09-845-416-34
; Sequence 34, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 4990
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-34

Query Match 93.7%; Score 2250; DB 13; Length 4990;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 141; Gaps 1;

QY 1 GGCAGTTCATGTGAGAGAGTGAAGTAACCTGGAACCTTATCAACAGCTTTAGAGAA 60
Db 1781 GGCAGTTCATGTGAGAGAGTGAAGTAACCTGGAACCTTATCAACAGCTTTAGAGAA 1840
QY 61 GTATTATCGTGGCTTCTTCTGTGAGAGACACATTGCAAGCAAGAGAGATTCTAAT 120
Db 1841 GTATTATCGTGGCTTCTTCTGTGAGAGACACATTGCAAGCAAGAGAGATTCTAAT 1900
QY 121 GATGTGAAGTGTGAAAGACAGTTTCTACTCATGAGGGGTACATGATGATTTGACA 180
Db 1901 GATGTGAAGTGTGAAAGACAGTTTCTACTCATGAGGGGTACATGATGATTTGACA 1960
QY 181 GCCCATCAGGGCCGGGTTGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGA 240
Db 1961 GCCCATCAGGGCCGGGTTGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGA 2020
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAGAGAGAGATGAATCTCCTTAATTCAAGA 300
Db 2021 AAATTATCAGAAGATGAAGAACTGAAGTACAGAGAGAGATGAATCTCCTTAATTCAAGA 2080
QY 301 TGGGAATGCTCAGGGTAGCTAGCATGGAACCAAGCAATTACATAGAGTTTAATG 360
Db 2081 TGGGAATGCTCAGGGTAGCTAGCATGGAACCAAGCAATTACATAGAGTTTAATG 2140
QY 361 GATCTCAGAATCAGAACTGAAGAGTGTGAATGACTGCTAACAAAAACAGAGAAAGA 420

Db 2141 GATCTCCAGAAATGAGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGA 2200
QY 421 ACAAGGAAAAATGAGAGAAAGAGCTCTTGAGACCTGATCTTGAAGACCTTAAACGCCAAGTA 480
Db 2201 ACAAGGAAAAATGAGAGAAAGAGCTCTTGAGACCTGATCTTGAAGACCTTAAACGCCAAGTA 2260
QY 481 CAACAACATAGGTGCTTCAAGAAAGATCTAGAACAGAAACAAGTCAGGCTCAATTCTCTC 540
Db 2261 CAACAACATAGGTGCTTCAAGAAAGATCTAGAACAGAAACAAGTCAGGCTCAATTCTCTC 2320
QY 541 ACTCACATGTTGGTGTAGTTGATGAATCTAGTGAGATCAGCAACTGCTTGGAA 600
Db 2321 ACTCACATGTTGGTGTAGTTGATGAATCTAGTGAGATCAGCAACTGCTTGGAA 2380
QY 601 GAACAACCTTAAGTATTTGGAGATCGATGGCAACAATCTGTAGATGAGAGAAAGACCGC 660
Db 2381 GAACAACCTTAAGTATTTGGAGATCGATGGCAACAATCTGTAGATGAGAGAAAGACCGC 2440
QY 661 TGGGTTCTTTTACAAGAC----- 678
Db 2441 TGGGTTCTTTTACAAGACCGCTGACCTAGCTCCTGGACTGACCATAATTGAGCCTCT 2500
QY 679 ----- 678
Db 2501 CCTACTCAGACTGTACTCTGGNGACACAACCTGTGTTACTAAGAAACTGCCATCTCC 2560
QY 679 -----ACTCATAGATTCTGCAACAG 699
Db 2561 AAACCTAGAAATGCCATCTTCTTGATGTGAGAGTACTACTCATAGATTACTGCAACAG 2620
QY 700 TTCCCCCTGGAACCTGGAAGTTTCTTGCTGGCTTACAGAAAGCTGAACAACACTGCCAAT 759
Db 2621 TTCCCCCTGGAACCTGGAAGTTTCTTGCTGGCTTACAGAAAGCTGAACAACACTGCCAAT 2680
QY 760 GTCTTACAGAGATGCTACCCGTAAGGAAAGGCTCCTAGAAAGACTCCAAAGGAGTAAAGAG 819
Db 2681 GTCTTACAGAGATGCTACCCGTAAGGAAAGGCTCCTAGAAAGACTCCAAAGGAGTAAAGAG 2740
QY 820 CTGATGAACAATGGCAAGACCTTCAAGGTGAATGAAGCTCACACAGATGTTATCAC 879
Db 2741 CTGATGAACAATGGCAAGACCTTCAAGGTGAATGAAGCTCACACAGATGTTATCAC 2800
QY 880 AACCTGATGAAAAACAGCCAAAAAATCTGAGATCCCTGGAAGGTTCCGATGATGCAGTC 939
Db 2801 AACCTGATGAAAAACAGCCAAAAAATCTGAGATCCCTGGAAGGTTCCGATGATGCAGTC 2860
QY 940 CTGTTACAAAGACGTTTGGATTAATGAATCTTCAAGTGAAGTGAATCTCGAAAAAGTCT 999
Db 2861 CTGTTACAAAGACGTTTGGATTAATGAATCTTCAAGTGAAGTGAATCTCGAAAAAGTCT 2920
QY 1000 CTCAACATTAGGTTCCCATTTGGAGCCAGTTCTGACAGTGAAGCGTCTGCACTTTCT 1059
Db 2921 CTCAACATTAGGTTCCCATTTGGAGCCAGTTCTGACAGTGAAGCGTCTGCACTTTCT 2980
QY 1060 CTGCAAGAACTTCTGGTGTGCTACAGCTGAAGATGAATTAAGCCGGCAGGCACT 1119
Db 2981 CTGCAAGAACTTCTGGTGTGCTACAGCTGAAGATGAATTAAGCCGGCAGGCACT 3040
QY 1120 ATTGAGGCGACTTTCCAGCAGTTCAAGAGCAGAACGATGTACATAGGGCTTCAAGAGG 1179
Db 3041 ATTGAGGCGACTTTCCAGCAGTTCAAGAGCAGAACGATGTACATAGGGCTTCAAGAGG 3100
QY 1180 GAATTGAATACTAAAGAACCTGTATCATGAGTACTCTGAGACTGTACGAATATTTCTG 1239
Db 3101 GAATTGAATACTAAAGAACCTGTATCATGAGTACTCTGAGACTGTACGAATATTTCTG 3160
QY 1240 ACAGAGAGCCTTTGGAAGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCTCCT 1299
Db 3161 ACAGAGAGCCTTTGGAAGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCTCCT 3220
QY 1300 GAGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAGAGCGCTGAGAGGTCAATACT 1359

Db 3221 GAGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAGAGCGGCTGAGAGGTCAATACT 3280
QY 1360 GAGTGGAAAAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTT 1419
Db 3281 GAGTGGAAAAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTT 3340
QY 1420 GAAAGACTCCAGAACTTCAAGAGCCACGGAAGCTGAGCTCAAGCTGCGCAAGCT 1479
Db 3341 GAAAGACTCCAGAACTTCAAGAGCCACGGAAGCTGAGCTCAAGCTGCGCAAGCT 3400
QY 1480 GAGGTGATCAAGGATCTTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCAAGAT 1539
Db 3401 GAGGTGATCAAGGATCTTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCAAGAT 3460
QY 1540 CACCTCGAGAAAGTCAAGGCACTTGAAGAGAAATTGCGCTCTGAAGAGAAAGCTGAGC 1599
Db 3461 CACCTCGAGAAAGTCAAGGCACTTGAAGAGAAATTGCGCTCTGAAGAGAAAGCTGAGC 3520
QY 1600 CACGTCAATGACCTTGTCTGCCAGCTTACCACTTTGGGCATTGAGCTCTCAACGTATAAC 1659
Db 3521 CACGTCAATGACCTTGTCTGCCAGCTTACCACTTTGGGCATTGAGCTCTCAACGTATAAC 3580
QY 1660 CTGACCACTCTGGAAGACTGAACACACAGATGGAAGCTTCTGACAGGTGGCGTGAAGAC 1719
Db 3581 CTGACCACTCTGGAAGACTGAACACACAGATGGAAGCTTCTGACAGGTGGCGTGAAGAC 3640
QY 1720 CGAGTCAAGCAGCTGCAATGAAGCCCAAGGACTTTGGTCCAGCATCTGACACTTTCTT 1779
Db 3641 CGAGTCAAGCAGCTGCAATGAAGCCCAAGGACTTTGGTCCAGCATCTGACACTTTCTT 3700
QY 1780 TCCAGCTGTGTCCAGGGTCCCTGGGAGAGAGGCCATCTGCCAAACAAGTGCCTACTAT 1839
Db 3701 TCCAGCTGTGTCCAGGGTCCCTGGGAGAGAGGCCATCTGCCAAACAAGTGCCTACTAT 3760
QY 1840 ATCAACCAAGAGACTCAACAACCTTGTGGGACCATCCCAAAATGACAGAGCTCTACCAG 1899
Db 3761 ATCAACCAAGAGACTCAACAACCTTGTGGGACCATCCCAAAATGACAGAGCTCTACCAG 3820
QY 1900 TCTTTAGCTGACCTGAATATGTCAGATTCTCAGCTTATAGAGCTGCCATGAAACTCCGA 1959
Db 3821 TCTTTAGCTGACCTGAATATGTCAGATTCTCAGCTTATAGAGCTGCCATGAAACTCCGA 3880
QY 1960 AGACTGCAAGAGCCCTTGTCTTGGATCTCTGAGCCTGTCAGCTGCATGTGATGCTTG 2019
Db 3881 AGACTGCAAGAGCCCTTGTCTTGGATCTCTGAGCCTGTCAGCTGCATGTGATGCTTG 3940
QY 2020 GACCAGCAACCTCAAGCAAAATGACAGCCCAATGATATCTGCAGATTAATAATTGT 2079
Db 3941 GACCAGCAACCTCAAGCAAAATGACAGCCCAATGATATCTGCAGATTAATAATTGT 4000
QY 2080 TTGACCACTATTATGACGCGCTGAGCAAGAGCAACAATTGTCAACGTCCTCTC 2139
Db 4001 TTGACCACTATTATGACGCGCTGAGCAAGAGCAACAATTGTCAACGTCCTCTC 4060
QY 2140 TGGGTGATATGTCTGAACTGGCTGTGAATGTTATGATAGGAGCAAGGAGG 2199
Db 4061 TGGGTGATATGTCTGAACTGGCTGTGAATGTTATGATAGGAGCAAGGAGG 4120
QY 2200 ATCCGTGTCCTGCTTTTAAAACTGGCATCATTTCCCTGTGTAAGACACATTTGGAAGAC 2259
Db 4121 ATCCGTGTCCTGCTTTTAAAACTGGCATCATTTCCCTGTGTAAGACACATTTGGAAGAC 4180
QY 2260 AAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACGAGCGCAGG 2319
Db 4181 AAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACGAGCGCAGG 4240
QY 2320 CTGGGCTCTCTTGTGATGATTTCTCAAAATTTCCAAAGACAGTTGGGTGAAGTGCATCC 2379
Db 4241 CTGGGCTCTCTTGTGATGATTTCTCAAAATTTCCAAAGACAGTTGGGTGAAGTGCATCC 4300
QY 2380 TTTGGGGCAGTAACATTGAGC 2401
Db 4301 TTTGGGGCAGTAACATTGAGC 4322

RESULT 8
US-09-845-416-2
; Sequence 2, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4182
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-2

Query Match 86.1%; Score 2067; DB 13; Length 4182;
Best Local Similarity 88.1%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 324; Gaps 1;

QY 1 GGCAGTTCATGATGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAAGAA 60
DB 1000 GGCAGTTCATGATGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAAGAA 1059
QY 61 GTATTATCGTGCTTCTTCTGCTGAGGACACATTCAGCAACAAGAGAGATTCTAAT 120
DB 1060 GTATTATCGTGCTTCTTCTGCTGAGGACACATTCAGCAACAAGAGAGATTCTAAT 1119
QY 121 GATGTGAAGTGTGAAGAGACAGTTTCATCTACTCATGAGGGGTACATGATGATTGACA 180
DB 1120 GATGTGAAGTGTGAAGAGACAGTTTCATCTACTCATGAGGGGTACATGATGATTGACA 1179
QY 181 GCCCATCAGGCGCGGTTGTAATATTTCTACATTTGGAAAGTAAGCTGATTGGAACAGGA 240
DB 1180 GCCCATCAGGCGCGGTTGTAATATTTCTACATTTGGAAAGTAAGCTGATTGGAACAGGA 1239
QY 241 AAATATCAGAAGATGAAGAACTGAAGTACAGAAGACAGATGAATCTCCTAAATTCAAGA 300
DB 1240 AAATATCAGAAGATGAAGAACTGAAGTACAGAAGACAGATGAATCTCCTAAATTCAAGA 1299
QY 301 TGGGAATGCTCAGGGTAGCTAGCATGGAAGAAACAAGCAATTTACATAGAGTTTAAATG 360
DB 1300 TGGGAATGCTCAGGGTAGCTAGCATGGAAGAAACAAGCAATTTACATAGAGTTTAAATG 1359
QY 361 GATCTCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGA 420
DB 1360 GATCTCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGA 1419
QY 421 ACAAGGAAATGAGGAAGAGCTCTTGACCTGATCTTGAAGACCTTAAACGCCAAGTA 480
DB 1420 ACAAGGAAATGAGGAAGAGCTCTTGACCTGATCTTGAAGACCTTAAACGCCAAGTA 1479
QY 481 CAACAACATTAAGGTGCTTCAAGAAGATCTAGAACAAGAAACAAGTCAGGGTCAATTCTCTC 540
DB 1480 CAACAACATTAAGGTGCTTCAAGAAGATCTAGAACAAGAAACAAGTCAGGGTCAATTCTCTC 1539
QY 541 ACTCATGTGTGTGTAGTGAATCTAGTGAAGATCAAGCAACTGCTGCTTTGGA 600
DB 1540 ACTCATGTGTGTGTAGTGAATCTAGTGAAGATCAAGCAACTGCTGCTTTGGA 1599
QY 601 GAACAACCTTAAGGTATTTGGAGATGATGGGCAACATCTGTAGATGACAGAAAGACCGC 660
DB 1600 GAACAACCTTAAGGTATTTGGAGATGATGGGCAACATCTGTAGATGACAGAAAGACCGC 1659
QY 661 TGGGTCTTTTACAAGAC----- 678

DB 1660 TGGGTCTTTTACAAGACATCCTTCTCAATGGCAACGTTCTTACTGAAGAACAGTGCCCTT 1719
QY 679 ----- 678
DB 1720 TTAGTGATGGCTTTCAGAAAAAGAGATGACGTGAACAGATTCAACAACACTGGCTTT 1779
QY 679 ----- 678
DB 1780 AAAGATCAAAATGAATGTATATCAAGTCTTCAAAAACTGGCGGTTTAAAAAGCGATCTA 1839
QY 679 ----- 678
DB 1840 GAAAAAGAAAGCAATCCATGGGCAACTGTATTCACTCAACAAGATCTTCTTCAACA 1899
QY 679 ----- 678
DB 1900 CTGAAGATTAAGTCAGTGACCCAGAAAGACGGAAGCATGGCTGATTAACCTTGGCCGGTGT 1959
QY 679 -----ACTCATAGATTACTGCAA 696
DB 1960 TGGGATAATTAGTCCAAAAACTGAAAGAGTACAGACAGACTCATAGATTACTGCAA 2019
QY 697 CAGTCCCGCTGACCTGGAAGATTTCTTGCCCTGGCTTACAGAAGCTGAACAACACTGCC 756
DB 2020 CAGTCCCGCTGACCTGGAAGATTTCTTGCCCTGGCTTACAGAAGCTGAACAACACTGCC 2079
QY 757 AATGTCTTACAGGATGCTACCCCTTAAGGAAAGGCTCCTTGAAGACTCCCAAGGAGTAA 816
DB 2080 AATGTCTTACAGGATGCTACCCCTTAAGGAAAGGCTCCTTGAAGACTCCCAAGGAGTAA 2139
QY 817 GAGCTGATGAACAATGGCAAGACCTCCAAGGTGAATTAAGCTCACACAGATGTTAT 876
DB 2140 GAGCTGATGAACAATGGCAAGACCTCCAAGGTGAATTAAGCTCACACAGATGTTAT 2199
QY 877 CACAACCTGGATGAAAAAGCAAAAAATCTCTGAGATCCCTGGAAGGTTCCGATGATGCA 936
DB 2200 CACAACCTGGATGAAAAAGCAAAAAATCTCTGAGATCCCTGGAAGGTTCCGATGATGCA 2259
QY 937 GTCCCTGTTACAAAGACGTTTGGATTAACATGAACCTTCAAGTGAAGTGAACCTTGGAAAAAG 996
DB 2260 GTCCCTGTTACAAAGACGTTTGGATTAACATGAACCTTCAAGTGAAGTGAACCTTGGAAAAAG 2319
QY 997 TCTCTCAACATTAGGTCCCATTTGGAAGCCAGTCTGACAGTGAAGCGTCTGCACTT 1056
DB 2320 TCTCTCAACATTAGGTCCCATTTGGAAGCCAGTCTGACAGTGAAGCGTCTGCACTT 2379
QY 1057 TCTCTGAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCA 1116
DB 2380 TCTCTGAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCA 2439
QY 1117 CCTATGAGAGCGCACTTTCAGCAGTTCAAGAGCAGAAAGATGTACATAGAGGCTTCAAG 1176
DB 2440 CCTATGAGAGCGCACTTTCAGCAGTTCAAGAGCAGAAAGATGTACATAGAGGCTTCAAG 2499
QY 1177 AGGGAATTGAAAACTAAAGAACTGTATCATGAGTACTCTTGAGACTGTAGAAATTTT 1236
DB 2500 AGGGAATTGAAAACTAAAGAACTGTATCATGAGTACTCTTGAGACTGTAGAAATTTT 2559
QY 1237 CTGACAGACAGCGCTTTGGAAGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCCCT 1296
DB 2560 CTGACAGACAGCGCTTTGGAAGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCCCT 2619
QY 1297 CCTGAGAGAGAGAGCCAGAGATGTCACTCGGCTTCTACGAAAGCAGAGCTGAAGAGTCAAT 1356
DB 2620 CCTGAGAGAGAGAGCCAGAGATGTCACTCGGCTTCTACGAAAGCAGAGCTGAAGAGTCAAT 2679
QY 1357 ACTGAGTGGGAAAAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAACC 1416
DB 2680 ACTGAGTGGGAAAAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAACC 2739
QY 1417 CTTGAAGACTCCAGGAACCTTCAAGAGGCCAGGATGAGCTGGAAGCTCAAGCTGGGCCAA 1476
DB 2740 CTTGAAGACTCCAGGAACCTTCAAGAGGCCAGGATGAGCTGGAAGCTCAAGCTGGGCCAA 2799

| | | | |
|----|------|--|------|
| QY | 1477 | GCTGAGGTGATCAAGGGATCTTGGCAGCCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAA | 1536 |
| Db | 2800 | GCTGAGGTGATCAAGGGATCTTGGCAGCCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAA | 2859 |
| QY | 1537 | GATCACCCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCCCTCTGAAAGAACGTG | 1596 |
| Db | 2860 | GATCACCCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCCCTCTGAAAGAACGTG | 2919 |
| QY | 1597 | AGCCACGTCAATGACTTGTCTCGCCAGCTTACCACTTTGGGCATTGAGCTCTCAACCGTAT | 1656 |
| Db | 2920 | AGCCACGTCAATGACTTGTCTCGCCAGCTTACCACTTTGGGCATTGAGCTCTCAACCGTAT | 2979 |
| QY | 1657 | AACCTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCGCTGAG | 1716 |
| Db | 2980 | AACCTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCGCTGAG | 3039 |
| QY | 1717 | GACCCAGTCAGGCAGCTGCATGAAAGCCACAGAGGACTTTGGTCCAGCATCTCAGCACTT | 1776 |
| Db | 3040 | GACCCAGTCAGGCAGCTGCATGAAAGCCACAGAGGACTTTGGTCCAGCATCTCAGCACTT | 3099 |
| QY | 1777 | CTTTCACGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTGCCAAACAAGTGCCCTAC | 1836 |
| Db | 3100 | CTTTCACGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTGCCAAACAAGTGCCCTAC | 3159 |
| QY | 1837 | TATATCAACCACGAGACTCAACAACACTTGCTGGGACCATCCCAAAATGACAGAGCTTAC | 1896 |
| Db | 3160 | TATATCAACCACGAGACTCAACAACACTTGCTGGGACCATCCCAAAATGACAGAGCTTAC | 3219 |
| QY | 1897 | CAGTCTTTAGCTGACCTGAAATAATGTCAAGATTCTCAGCTTATAGACTGCCATGAACTC | 1956 |
| Db | 3220 | CAGTCTTTAGCTGACCTGAAATAATGTCAAGATTCTCAGCTTATAGACTGCCATGAACTC | 3279 |
| QY | 1957 | CGAAGACTGCAGAAGGCCCTTGCTTGATCTCTTGAGCCTGTCACTGCATGTGATGCC | 2016 |
| Db | 3280 | CGAAGACTGCAGAAGGCCCTTGCTTGATCTCTTGAGCCTGTCACTGCATGTGATGCC | 3339 |
| QY | 2017 | TTGGAACGACACAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTTAAT | 2076 |
| Db | 3340 | TTGGAACGACACAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTTAAT | 3399 |
| QY | 2077 | TGTTTGACCACCTATTTATGACCGCCTCGAGCAAGAGCAACAATTGGTCAACGTCCCT | 2136 |
| Db | 3400 | TGTTTGACCACCTATTTATGACCGCCTCGAGCAAGAGCAACAATTGGTCAACGTCCCT | 3459 |
| QY | 2137 | CTCTGCCGTGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGG | 2196 |
| Db | 3460 | CTCTGCCGTGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGG | 3519 |
| QY | 2197 | AGGATCCGTGTCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAA | 2256 |
| Db | 3520 | AGGATCCGTGTCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAA | 3579 |
| QY | 2257 | GACCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAAGGC | 2316 |
| Db | 3580 | GACCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTGTGACCAAGGC | 3639 |
| QY | 2317 | AGGCTGGGCTCTCTTCTGCATGATTTCTATCCAAATTCACAGACAGTTGGGTGAAGTTGCA | 2376 |
| Db | 3640 | AGGCTGGGCTCTCTTCTGCATGATTTCTATCCAAATTCACAGACAGTTGGGTGAAGTTGCA | 3699 |
| QY | 2377 | TCCTTTGGGGGCGAGTAACATTGAGC 2401 | |
| Db | 3700 | TCCTTTGGGGGCGAGTAACATTGAGC 3724 | |

RESULT 9
US-09-845-416-27
; Sequence 27, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE

```

; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 5149
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-27

```

| | | | | |
|-----------------------|--------------|--|------------|--------------|
| Query Match | 86.1%; | Score 2067; | DB 13; | Length 5149; |
| Best Local Similarity | 89.1%; | Pred. No. 0; | | |
| Matches 2401; | Conservative | 0; | Mismatches | 0; |
| | | | Indels | 324; |
| | | | Gaps | 1; |
| QY | 1 | GGCAGTTCAATGATGAGAGAGTGAAGTAAACCTGGAACCGTTATCAACAGCTTTAGAAAGAA | 60 | |
| DB | 1757 | GGCAGTTCAATGATGAGAGAGTGAAGTAAACCTGGAACCGTTATCAACAGCTTTAGAAAGAA | 1816 | |
| QY | 61 | GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTTGCAAGCACCAAGAGAGATTCTAAT | 120 | |
| DB | 1817 | GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTTGCAAGCACCAAGAGAGATTCTAAT | 1876 | |
| QY | 121 | GATGTGAAGTGTGAAGAAGACCAGTTTCATACCTCATGAGGGGTACATGATGAAATTGACA | 180 | |
| DB | 1877 | GATGTGAAGTGTGAAGAAGACCAGTTTCATACCTCATGAGGGGTACATGATGAAATTGACA | 1936 | |
| QY | 181 | GCCCATCAGGGCCGGGTTGCTAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGA | 240 | |
| DB | 1937 | GCCCATCAGGGCCGGGTTGCTAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGA | 1996 | |
| QY | 241 | AAATTATCAGAAAGATGAAGAAACTGAAGTACAAAGACAGATGAATCTCCTAAATTCAGA | 300 | |
| DB | 1997 | AAATTATCAGAAAGATGAAGAAACTGAAGTACAAAGACAGATGAATCTCCTAAATTCAGA | 2056 | |
| QY | 301 | TGGGAATGCTCTCAGGGTAGCTAGCATGGAACAAAGCAATTTACATAGAGTTTAAATG | 360 | |
| DB | 2057 | TGGGAATGCTCTCAGGGTAGCTAGCATGGAACAAAGCAATTTACATAGAGTTTAAATG | 2116 | |
| QY | 361 | GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGCTAACCAAAAACAGAAAGAA | 420 | |
| DB | 2117 | GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGCTAACCAAAAACAGAAAGAA | 2176 | |
| QY | 421 | ACAAGGAAATGAGAGAGAGCCCTTGGACCTGATCTTGAAGACCTAAACCGCCAAGTA | 480 | |
| DB | 2177 | ACAAGGAAATGAGAGAGAGCCCTTGGACCTGATCTTGAAGACCTAAACCGCCAAGTA | 2236 | |
| QY | 481 | CAACAACATAAGGTGCTTCAAGAGATCTAGAACAAAGCAAGTCAGGGTCAATTCTCTC | 540 | |
| DB | 2237 | CAACAACATAAGGTGCTTCAAGAGATCTAGAACAAAGCAAGTCAGGGTCAATTCTCTC | 2296 | |
| QY | 541 | ACTCACATGTTGTTGTTAGTTGATGAATCTTAGTGAGATCACGCCAATGCTGCTTTGGAA | 600 | |
| DB | 2297 | ACTCACATGTTGTTGTTAGTTGATGAATCTTAGTGAGATCACGCCAATGCTGCTTTGGAA | 2356 | |
| QY | 601 | GAACAACCTTAAGTAATGGAGATCGATGGCAACAATCTGTAGATGACAGAAAGACCGC | 660 | |
| DB | 2357 | GAACAACCTTAAGTAATGGAGATCGATGGCAACAATCTGTAGATGACAGAAAGACCGC | 2416 | |
| QY | 661 | TGGGTTCTTTACAAGAC----- | 678 | |
| DB | 2417 | TGGGTTCTTTACAAGACATCCTTCTCAAATGGCAAGTCTTACTGAAGAACAGTGCCTT | 2476 | |
| QY | 679 | ----- | 678 | |
| DB | 2477 | TTTAGTCATGGCTTTCAGAAAAAAGAGATGCAGTGAACAAGATTCAACACACTGGCTTT | 2536 | |
| QY | 679 | ----- | 678 | |
| DB | 2537 | AAAGATCAAAATGAATGTTATCAAGTCTTCAAAAACTGGCCGTTTAAAGCCGATCTA | 2596 | |

QY 679 ----- 678
Db 2597 GAAAAAGAAAGCAATCCATGGGCAAACTGTATTCACTCAACAAGATCTTTTCAACA 2656
QY 679 ----- 678
Db 2657 CTGAAGATAAGTCAGTGACCAGAGAAGCGAAGCATGGCTGATACTTTGCCGGTGT 2716
QY 679 ----- ACTCATAGATTACTGCAA 696
Db 2717 TGGGATAATTAGTCCAAAACTGAAAGAGTACAGCAGACTCATAGATTACTGCAA 2776
QY 697 CAGTCCCOCTGGACCTGSAAGTTTCTTGCCCTGCTTACAGAAGCTGAACAACCTGCC 756
Db 2777 CAGTCCCOCTGGACCTGSAAGTTTCTTGCCCTGCTTACAGAAGCTGAACAACCTGCC 2836
QY 757 AATGTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAGAAGCTCCAGAGGAGTAAAA 816
Db 2837 AATGTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAGAAGCTCCAGAGGAGTAAAA 2896
QY 817 GAGCTGATGAACAATGGCAAGACTCCAAAGGTGAATTTGAAGCTCACACAGATGTTAT 876
Db 2897 GAGCTGATGAACAATGGCAAGACTCCAAAGGTGAATTTGAAGCTCACACAGATGTTAT 2956
QY 877 CACAACCTGGATGAAAAAGCAAGCAAAAAATCCTGAGATCCTGGAAGGTTCCGATGATGCA 936
Db 2957 CACAACCTGGATGAAAAAGCAAGCAAAAAATCCTGAGATCCTGGAAGGTTCCGATGATGCA 3016
QY 937 GTCTGTGTTACAAAGACGTTTGGATTAACATGAACCTCAAGTGAAGTGAACCTCGAAAAAG 996
Db 3017 GTCTGTGTTACAAAGACGTTTGGATTAACATGAACCTCAAGTGAAGTGAACCTCGAAAAAG 3076
QY 997 TCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCACTGGAAGCGTCTGCACTT 1056
Db 3077 TCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCACTGGAAGCGTCTGCACTT 3136
QY 1057 TCTCTGAGAGAACTTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAAGCCGAGGCA 1116
Db 3137 TCTCTGAGAGAACTTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAAGCCGAGGCA 3196
QY 1117 CCTATTGAGGCGACTTTCCAGAGTTCAAGAGCAGAAAGATGTACATAGGCGCTTCAAG 1176
Db 3197 CCTATTGAGGCGACTTTCCAGAGTTCAAGAGCAGAAAGATGTACATAGGCGCTTCAAG 3256
QY 1177 AGGGAATTGAAAACTAAGAAGACTGTAATCATGAGTACTTTGAGACTGTACGAATATT 1236
Db 3257 AGGGAATTGAAAACTAAGAAGACTGTAATCATGAGTACTTTGAGACTGTACGAATATT 3316
QY 1237 CTGACAGAGCAGCCTTTGGAGAGCACTAGAAAACTCTACCAAGAGCCCAAGAGCTGCC 1296
Db 3317 CTGACAGAGCAGCCTTTGGAGAGCACTAGAAAACTCTACCAAGAGCCCAAGAGCTGCC 3376
QY 1297 CCTGAGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCAT 1356
Db 3377 CCTGAGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCAT 3436
QY 1357 ACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACC 1416
Db 3437 ACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACC 3496
QY 1417 CTTGAAAGACTCCAGGAACCTTCAAGAGGCCAGGATGAGCTGGAACCTCAAGCTGCGCAA 1476
Db 3497 CTTGAAAGACTCCAGGAACCTTCAAGAGGCCAGGATGAGCTGGAACCTCAAGCTGCGCAA 3556
QY 1477 GCTGAGGTGATCAAGGATCTGCGCAGCCCGTGGCGATCTCTCATTTGACTCTCTCAA 1536
Db 3557 GCTGAGGTGATCAAGGATCTGCGCAGCCCGTGGCGATCTCTCATTTGACTCTCTCAA 3616
QY 1537 GATCACTCGAGAAAGTCAAGGCACTTGCAGGAGAAATTGCGCTCTGAAAGAGAACGTG 1596
Db 3617 GATCACTCGAGAAAGTCAAGGCACTTGCAGGAGAAATTGCGCTCTGAAAGAGAACGTG 3676

QY 1597 AGCCAGTCAATGACCTTGTCTGCGCCAGCTTACCACCTTTGGGCATTACGCTCTACCCGTAT 1656
Db 3677 AGCCAGTCAATGACCTTGTCTGCGCCAGCTTACCACCTTTGGGCATTACGCTCTACCCGTAT 3736
QY 1657 AACCTCAGCACTCTGGAAGACTGAAACACCAAGATGAAAGCTTCTGAGGTGGCGGTGAG 1716
Db 3737 AACCTCAGCACTCTGGAAGACTGAAACACCAAGATGAAAGCTTCTGAGGTGGCGGTGAG 3796
QY 1717 GACCGAGTCAGGCAAGCTGCATGAAAGCCCAAGGAGCTTTGGTCCAGCATCTCAGCACTTT 1776
Db 3797 GACCGAGTCAGGCAAGCTGCATGAAAGCCCAAGGAGCTTTGGTCCAGCATCTCAGCACTTT 3856
QY 1777 CTTTCCAGCTCTGTCCAGGGTCCCTGGAGAGAGGCCATCTCGCCAAACAAAGTGCCTAC 1836
Db 3857 CTTTCCAGCTCTGTCCAGGGTCCCTGGAGAGAGGCCATCTCGCCAAACAAAGTGCCTAC 3916
QY 1837 TATATCAACCAAGAGACTCAACAACTTGCTGGGACCAATCCCAAAATGACAGAGCTCTAC 1896
Db 3917 TATATCAACCAAGAGACTCAACAACTTGCTGGGACCAATCCCAAAATGACAGAGCTCTAC 3976
QY 1897 CAGTCTTTAGCTGACCTGAATTAATGTACAGATTTCTAGCTTATAGAGCTGCCATGAACCTC 1956
Db 3977 CAGTCTTTAGCTGACCTGAATTAATGTACAGATTTCTAGCTTATAGAGCTGCCATGAACCTC 4036
QY 1957 CGAAGACTGAGAGAGGCCCTTTGCTTGATCTCTTGAAGCTGTACAGCTGCATGTGATGCC 2016
Db 4037 CGAAGACTGAGAGAGGCCCTTTGCTTGATCTCTTGAAGCTGTACAGCTGCATGTGATGCC 4096
QY 2017 TTGACACGACACAACCTCAAGCAAAATGACAGGCCCATGGATATCCTGCAGATTAATTAAT 2076
Db 4097 TTGACACGACACAACCTCAAGCAAAATGACAGGCCCATGGATATCCTGCAGATTAATTAAT 4156
QY 2077 TGTTTGACCACTATTTATGACCGCCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCCT 2136
Db 4157 TGTTTGACCACTATTTATGACCGCCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCCT 4216
QY 2137 CTCTGCGTGGATATGTGTCTGAACTGGCTGTGTAATGTTATGATACGGGACCAAGAGG 2196
Db 4217 CTCTGCGTGGATATGTGTCTGAACTGGCTGTGTAATGTTATGATACGGGACCAAGAGG 4276
QY 2197 AGGATCCGTGCTCTGCTTTTAAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGA 2256
Db 4277 AGGATCCGTGCTCTGCTTTTAAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGA 4336
QY 2257 GACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTGTGACAGCGC 2316
Db 4337 GACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTGTGACAGCGC 4396
QY 2317 AGGCTGGGCTCTCTCTGCAATGATTTATCCAAATTCAGAGAGAGTGGGTGAAGTGA 2376
Db 4397 AGGCTGGGCTCTCTCTGCAATGATTTATCCAAATTCAGAGAGAGTGGGTGAAGTGA 4456
QY 2377 TCCTTTGGGGCAGTAACATTGAGC 2401
Db 4457 TCCTTTGGGGCAGTAACATTGAGC 4481

RESULT 10
US-10-149-736-41
; Sequence 41, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: US-06968
; CURRENT APPLICATION NUMBER: US/10/149, 736
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238, 848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96

SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; LENGTH: 5462
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-41

Query Match 76.6%; Score 1839.4; DB 13; Length 5462;
Best Local Similarity 88.1%; Pred. No. 0;
Matches 2115; Conservative 0; Mismatches 76; Indels 210; Gaps 2;

QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 60
Db 1199 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 1258
QY 61 GTATTATCGTGGCTTCTTCTGCTGAGGACATTGCAAGCAGAGAGAGATTCTTAAT 120
Db 1259 GTATTATCGTGGCTTCTTCTGCTGAGGACATTGCAAGCAGAGAGATTCTTAAT 1318
QY 121 GATGTGAAGTGGTGAAGAACCCAGTTTCATCTCATGAGGGGTACATGATGATTGACA 180
Db 1319 GATGTGAAGTGGTGAAGAACCCAGTTTCATCTCATGAGGGGTACATGATGATTGACA 1378
QY 181 GCCCATCAGGGCCGGGTGGTATATTTCTCAATTGGAGTAAGCTGATTGAAACAGGA 240
Db 1379 GCCCATCAGGGCCGGGTGGTATATTTCTCAATTGGAGTAAGCTGATTGAAACAGGA 1438
QY 241 AAATTATCAGAGATGAAGAACTGAAGTCAAGAGCAGATGAATCTCTAAATTGACAA 300
Db 1439 AAATTATCAGAGATGAAGAACTGAAGTCAAGAGCAGATGAATCTCTAAATTGACAA 1498
QY 301 TGGGAATGCTCAGGGGTAGCTAGCATGGAAGAAACAAGCAATTTACATAGATTTTATG 360
Db 1499 TGGGAATGCTCAGGGGTAGCTAGCATGGAAGAAACAAGCAATTTACATAGATTG 1558
QY 361 GATCTCCAGATCAGAAACTGAAGAAGTTGAATGACTGGCTAACAAAAACAGAGAAAGA 420
Db 1559 ACCA----- 1562
QY 421 ACAAGGAAATGAGAGAGAGCCCTTTGGACCTGATCTTGAAGACCTAAACGCCAGTA 480
Db 1563 -----CTATTGGAGCCTCTCTCTACTCAGACTGTACTCTGTG 1600
QY 481 CAACAACATAAGTGTCTCAAGAAGATCTGAACAAGAACAGAGTCAAGGTCATTTCTTC 540
Db 1601 ACACAACCTGTGTTACTAAGGAACTGCCATCTCCAAACTAGAAATGCCATCTTCTTG 1660
QY 541 ACTCACATGGTGTGTAGTTGATGAATCTAGTGAGATCAGCGCACTGCTGCTTTGGA 600
Db 1661 ATGTTGAG----- 1669
QY 601 GAACAACCTTAAGGATTGGGAGATCGATGGGCAACATCTGTAGATGGAACAGAGACCGC 660
Db 1670 ----- 1669
QY 661 TGGGTTCTTTACAAGACACTCATAGATTACTGCAACAGTTCCCTGGACCTGAAAG 720
Db 1670 -----CATAGATTACTGCAACAGTTCCCTGGACCTGAAAG 1708
QY 721 TTTCTTGCTGCTTACAGAAGCTGAAGCACTGCCAATGTCTTACAGATGCTACCCGT 780
Db 1709 TTTCTTGCTGCTTACAGAAGCTGAAGCACTGCCAATGTCTTACAGATGCTACCCGT 1768
QY 781 AAGGAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAAGAGCTGATGAACAATGGCAAGC 840
Db 1769 AAGGAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAAGAGCTGATGAACAATGGCAAGC 1828
QY 841 CTCCAAGGTGAATGAAGCTCACACAGATGTTTATCACAACCTGGATGAAGAACAGCCAA 900
Db 1829 CTCCAAGGTGAATGAAGCTCACACAGATGTTTATCACAACCTGGATGAAGAACAGCCAA 1888

QY 901 AAATCCTGAGATCCCTGGAGGTTCCGATGATGACAGTCTCTTTACAAGACGTTGGAT 960
Db 1889 AAATCCTGAGATCCCTGGAGGTTCCGATGATGACAGTCTCTCTTTACAAGACGTTGGAT 1948
QY 961 AACATGAATCTCAAGTGAAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAGTCCCATTTG 1020
Db 1949 AACATGAATCTCAAGTGAAGTGAACCTTCGGAAGAAAGTCTCTCTCAACATTAGTCCCATTTG 2008
QY 1021 GAAGCAGTTCGACCCAGTGAAGGCTCTGACACCTTCTCTGACAGGAACCTCTGTGTGG 1080
Db 2009 GAAGCAGTTCGACCCAGTGAAGGCTCTGACACCTTCTCTGACAGGAACCTCTGTGTGG 2068
QY 1081 CTACAGCTGAAGATGATGAATTAAGCCGCGACAGGACCTTATTTGAAGGCTTTCCAGCA 1140
Db 2069 CTACAGCTGAAGATGATGAATTAAGCCGCGACAGGACCTTATTTGAAGGCTTTCCAGCA 2128
QY 1141 GTTCAGAGCAGAAAGATGTACATAGGGCTTCAAGAGGGAATTGAAAACCTAAAGAACT 1200
Db 2129 GTTCAGAGCAGAAAGATGTACATAGGGCTTCAAGAGGGAATTGAAAACCTAAAGAACT 2188
QY 1201 GTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGACAGCCTTTGGAAGGA 1260
Db 2189 GTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGACAGCCTTTGGAAGGA 2248
QY 1261 CTAGAGAACTCTACCAGAGAGCCGAGAGAGCTGCTCTCTGAGAGAGAGCCAGAAATGTC 1320
Db 2249 CTAGAGAACTCTACCAGAGAGCCGAGAGAGCTGCTCTCTGAGAGAGAGCCAGAAATGTC 2308
QY 1321 ACTGGCTTCTACGAAGAGGCTGAGAGGTCAATATCTAGTGGGAAAAATTGAACCTG 1380
Db 2309 ACTGGCTTCTACGAAGAGGCTGAGAGGTCAATATCTAGTGGGAAAAATTGAACCTG 2368
QY 1381 CACTCCGCTGACTGGCAGAGAAATTAATGATGAGACCTTTGAAAGACTCCAGGAACCTTCA 1440
Db 2369 CACTCCGCTGACTGGCAGAGAAATTAATGATGAGACCTTTGAAAGACTCCAGGAACCTTCA 2428
QY 1441 GAGGCCACGGATGAGCTGGAACCTCAAGCTGCGCCAAAGCTGAGTGATCAAGGGATCCTGG 1500
Db 2429 GAGGCCACGGATGAGCTGGAACCTCAAGCTGCGCCAAAGCTGAGTGATCAAGGGATCCTGG 2488
QY 1501 CAGCCCGTGGCGATCTCTCATTTGACTCTCTCCAGATCACCCTGAGAAAGTCAAGGCA 1560
Db 2489 CAGCCCGTGGCGATCTCTCATTTGACTCTCTCCAGATCACCCTGAGAAAGTCAAGGCA 2548
QY 1561 CTTGAGAGAGAAATGCGCCTCTGAAGAGAAAGTGAAGCCAGTCAATGACCTTGCTGCG 1620
Db 2549 CTTGAGAGAGAAATGCGCCTCTGAAGAGAAAGTGAAGCCAGTCAATGACCTTGCTGCG 2608
QY 1621 CAGCTTACCACTTTGGGCATTCACTCTCAACGCTATTAACCTCAGCACTCTGGAAGACCTG 1680
Db 2609 CAGCTTACCACTTTGGGCATTCACTCTCAACGCTATTAACCTCAGCACTCTGGAAGACCTG 2668
QY 1681 AACACCAAGTGAAGCTTCTGACAGTGGCCGTGAGAGACCGAGTCAAGGAGCTGCATGAA 1740
Db 2669 AACACCAAGTGAAGCTTCTGACAGTGGCCGTGAGAGACCGAGTCAAGGAGCTGCATGAA 2728
QY 1741 GCCCAGAGGACTTTGTCAGACATCTCAGCACTTTCTTCCAGTCTGTCCAGGGTCCC 1800
Db 2729 GCCCAGAGGACTTTGTCAGACATCTCAGCACTTTCTTCCAGTCTGTCCAGGGTCCC 2788
QY 1801 TGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCAAGAGACTCAACA 1860
Db 2789 TGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCAAGAGACTCAACA 2848
QY 1861 ACTTGCTGGAGCCATCCCAAAATGACAGAGCTTACCAAGTCTTTAGCTGACCTGAATAAT 1920
Db 2849 ACTTGCTGGAGCCATCCCAAAATGACAGAGCTTACCAAGTCTTTAGCTGACCTGAATAAT 2908
QY 1921 GTCAGATTCTCAGCTTATAGGACTGCGCATGAAACTCCGAAGACTGCAAGAGGCCCTTTGC 1980
Db 2909 GTCAGATTCTCAGCTTATAGGACTGCGCATGAAACTCCGAAGACTGCAAGAGGCCCTTTGC 2968
QY 1981 TTGATCTCTTGAGCCTGTCAAGTGCATGTGATGCCCTTGGAACCAAGCAACTCAAGCAA 2040

```
Db      2969 TTGATCTCTTGAGCCTGTGACGTGATGATGCTTGACCAGACAACCTCAAGCAA 3028
QY      2041 AATGACCAAGCCCATGATATCCTGCAGATTATTAATTGTTGACCACTATTTAAGCCGC 2100
Db      3029 AATGACCAAGCCCATGATATCCTGCAGATTATTAATTGTTGACCACTATTTAAGCCGC 3088
QY      2101 CTGAGCAGAGACACAACAATTGTCACGTCCTCTGCGTGGATATGTGTGAAC 2160
Db      3089 CTGAGCAGAGACACAACAATTGTCACGTCCTCTGCGTGGATATGTGTGAAC 3148
QY      2161 TGGCTGTGAATGTTTATGATACGGAGCAAGAGGATCGGTCTGCTTTTAA 2220
Db      3149 TGGCTGTGAATGTTTATGATACGGAGCAAGAGGATCGGTCTGCTTTTAA 3208
QY      2221 ACTGCATCATTTCCCTGTGTAAAGCATTTGGAAGACAGTACATACCTTTTCAAG 2280
Db      3209 ACTGCATCATTTCCCTGTGTAAAGCATTTGGAAGACAGTACATACCTTTTCAAG 3268
QY      2281 CAAGTGGCAAGTTCAACAGGATTTTGAACCGCAGCGCTGGGCTCTCTTGCATGAT 2340
Db      3269 CAAGTGGCAAGTTCAACAGGATTTTGAACCGCAGCGCTGGGCTCTCTTGCATGAT 3328
QY      2341 TCTATCCAAATTCACAGACAGTTGGTGAAGTTGCATCCTTTGGGGCAGTAACATTGAG 2400
Db      3329 TCTATCCAAATTCACAGACAGTTGGTGAAGTTGCATCCTTTGGGGCAGTAACATTGAG 3388
QY      2401 C 2401
Db      3389 C 3389
```

```
RESULT 11
US-09-845-416-10
; Sequence 10, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: D31142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 3531
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-10
```

```
Query Match      72.3%; Score 1737; DB 13; Length 3531;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 1765; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY      576 AGATCAGCACTGCTGCTTGGAGAGACAACCTTAAGGTATGGAGATCGATGGGCAAA 635
Db      1248 AGAGATGAGAAACTGAAGTACAGAGCAGATGAATCTCTAAATTCAAGATGGGAATG 1307
QY      636 CATCTGTAGATGACAGAGAACCGCTGGGTTCTTTTACAAGACACTCATAGATTACTGCA 695
Db      1308 CCTCAGGGTAGCTAGCATGGAATAACAAGCAATTACATAGAACTCATAGATTACTGCA 1367
QY      696 ACAGTTCCTCCCTGACCTGGAATAAGTTTCTTGCTGGCTTACAGAAAGCTGAACAACCTGC 755
Db      1368 ACAGTTCCTCCCTGACCTGGAATAAGTTTCTTGCTGGCTTACAGAAAGCTGAACAACCTGC 1427
QY      756 CAATGTCTTACAGGATGTACCCCGTAAGAAAGGCTCTTAGAAGACTCCAAGGAGATAAA 815
Db      1428 CAATGTCTTACAGGATGTACCCCGTAAGAAAGGCTCTTAGAAGACTCCAAGGAGATAAA 1487
```

```
QY      816 AGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAATTTGAAGCTCACACAGATGTTTA 875
Db      1488 AGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAATTTGAAGCTCACACAGATGTTTA 1547
QY      876 TCACAACTTGATGAAACACAGCCAAATAATCCTGAGATCCCTGGAAAGSTTCCGATGATGC 935
Db      1548 TCACAACTTGATGAAACACAGCCAAATAATCCTGAGATCCCTGGAAAGSTTCCGATGATGC 1607
QY      936 AGTCCTGTACAAAGACGTTTGGATACATGAACCTTCAAGTGGAGTGAACCTTGGAAAAA 995
Db      1608 AGTCCTGTACAAAGACGTTTGGATACATGAACCTTCAAGTGGAGTGAACCTTGGAAAAA 1667
QY      996 GTCTCTCAACATTAAGTCCCATTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCT 1055
Db      1668 GTCTCTCAACATTAAGTCCCATTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCT 1727
QY      1056 TTCTCTCAGAACTTCTGTGTGCTACAGCTGAAGATGAATTAAGCCGCGCAGGC 1115
Db      1728 TTCTCTCAGAACTTCTGTGTGCTACAGCTGAAGATGAATTAAGCCGCGCAGGC 1787
QY      1116 ACCTATTGAGCGCACTTTCAGAGTTCAAGAGCAGACGAATGATAGAGGCTTCAA 1175
Db      1788 ACCTATTGAGCGCACTTTCAGAGTTCAAGAGCAGACGAATGATAGAGGCTTCAA 1847
QY      1176 GAGGAATTGAAAACTAAAGAACTGTATCATGAGTACTCTTGAGACTGTACGAATATT 1235
Db      1848 GAGGAATTGAAAACTAAAGAACTGTATCATGAGTACTCTTGAGACTGTACGAATATT 1907
QY      1236 TCTGACAGACAGCCTTTGGAAGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCC 1295
Db      1908 TCTGACAGACAGCCTTTGGAAGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCC 1967
QY      1296 TCCTGAGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAA 1355
Db      1968 TCCTGAGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAA 2027
QY      1356 TACTGAGTGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGAC 1415
Db      2028 TACTGAGTGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGAC 2087
QY      1416 CCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGAGTGAAGCTCAAGCTGCCCA 1475
Db      2088 CCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGAGTGAAGCTCAAGCTGCCCA 2147
QY      1476 AGCTGAGTGATCAAGGGATCCTGCGACGCCGCTGGCGGATCTCTCATTTGACTCTTCCA 1535
Db      2148 AGCTGAGTGATCAAGGGATCCTGCGACGCCGCTGGCGGATCTCTCATTTGACTCTTCCA 2207
QY      1536 AGATCACCTCGAAAAAGTCAAGGCACTTCGAGAGAGAAATGCGCTCGAAAGAGAAAGCT 1595
Db      2208 AGATCACCTCGAAAAAGTCAAGGCACTTCGAGAGAGAAATGCGCTCGAAAGAGAAAGCT 2267
QY      1596 GAGCCACGTCAATGACCTTGCTGCGCACTTACCACTTTGGGCAATTCAAGCTCTCACCGTA 1655
Db      2268 GAGCCACGTCAATGACCTTGCTGCGCACTTACCACTTTGGGCAATTCAAGCTCTCACCGTA 2327
QY      1656 TAACCTCAGCACTCTGGAAGACTGAAACACAGATGGAAGCTTCTGCAAGTGGCGCTGCA 1715
Db      2328 TAACCTCAGCACTCTGGAAGACTGAAACACAGATGGAAGCTTCTGCAAGTGGCGCTGCA 2387
QY      1716 GGACCGAGTCAGGACGTGATGAAGCCACAGAGGACTTTGGTCCAGCATCTCAGCACTT 1775
Db      2388 GGACCGAGTCAGGACGTGATGAAGCCACAGAGGACTTTGGTCCAGCATCTCAGCACTT 2447
QY      1776 TCTTTCAAGTCTGTCCAGGGTCCCTGGAGAGAGCACTCTGCCAAACAAGTGCCCTA 1835
Db      2448 TCTTTCAAGTCTGTCCAGGGTCCCTGGAGAGAGCACTCTGCCAAACAAGTGCCCTA 2507
QY      1836 CTATATCAACCAAGAGACTCAAACTTGTCTGGACCAATCCCAAAATGACAGAGCTCTA 1895
Db      2508 CTATATCAACCAAGAGACTCAAACTTGTCTGGACCAATCCCAAAATGACAGAGCTCTA 2567
QY      1896 CCAAGTCTTAGCTGACCTGAATAATGTCAAGATTCTCAGCTTATAGAGACTGCCATGAACCT 1955
```

Db 2568 CCAAGTCTTAGCTGACCTGATAATATGTAGATTTCTCAGCTTATAGAGCTGCCATGAACCT 2627
QY 1956 CCGAAGACTGCAGAAAGCCCTTGTCTTGATCTCTTGAGCCTGTACGTGCATGTGATGC 2015
Db 2628 CCGAAGACTGCAGAAAGCCCTTGTCTTGATCTCTTGAGCCTGTACGTGCATGTGATGC 2687
QY 2016 CTTGGACCAAGCACAACCTTCAAGCAAAATGACCAAGCCCATGGATATCTCGAGATTATTA 2075
Db 2688 CTTGGACCAAGCACAACCTTCAAGCAAAATGACCAAGCCCATGGATATCTCGAGATTATTA 2747
QY 2076 TTGTTTGACCACTATTATGACCCGCTTGAGCAGACACACAATTTGGTCAACGTCCC 2135
Db 2748 TTGTTTGACCACTATTATGACCCGCTTGAGCAGACACACAATTTGGTCAACGTCCC 2807
QY 2136 TCTCTGCGTGATATGTGTCTGAACCTGGCTGCTGAATGTTATGATACGGAGCAGACAGG 2195
Db 2808 TCTCTGCGTGATATGTGTCTGAACCTGGCTGCTGAATGTTATGATACGGAGCAGACAGG 2867
QY 2196 GAGGATCCGTGCTGCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGACACATTTGGA 2255
Db 2868 GAGGATCCGTGCTGCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGACACATTTGGA 2927
QY 2256 AGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAAGC 2315
Db 2928 AGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAAGC 2987
QY 2316 CAGGCTGGGCTCTCTTCTGATGATTTCTATCCAAATTCACAGACAGTGGGTGAAGTGC 2375
Db 2988 CAGGCTGGGCTCTCTTCTGATGATTTCTATCCAAATTCACAGACAGTGGGTGAAGTGC 3047
QY 2376 ATCCTTTGGGGGAGTAACATTGAGC 2401
Db 3048 ATCCTTTGGGGGAGTAACATTGAGC 3073

RESULT 12
US-09-845-416-30
; Sequence 30, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 4498
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-30

Query Match 72.3%; Score 1737; DB 13; Length 4498;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 1765; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 576 AGATCAGCAACTGCTGCTTTGGAAGAACAACTTAAGTATTGGAGATCGATGGGCAAA 635
Db 2005 AGAAGATGAAGAACTGAAGTACAAGACAGATGAATCTCTAAATTCAGATGGGAATG 2064
QY 636 CATCTGTAGATGACAGAGACCGCTGGGTTCTTTACAAGACACTCATAGATTACTGCA 695
Db 2065 CCTCAGGGTAGCTAGCATGMAAAACAAAGCAATTTACATAGAACTCATAGATTACTGCA 2124
QY 696 ACAATTCCCTGACCTGMAAAGTTTCTTGCTGCTTACAGAGCTGAACAACACTGC 755
Db 2125 ACAATTCCCTGACCTGMAAAGTTTCTTGCTGCTTACAGAGCTGAACAACACTGC 2184

QY 756 CAATGCTCTACAGATGCTAACCCGTAAAGAAAGCCTCCTAGAGACTTCCAGGAGTAA 815
Db 2185 CAATGCTCTACAGATGCTAACCCGTAAAGAAAGCCTCCTAGAGACTTCCAGGAGTAA 2244
QY 816 AGAGCTGATGAACAAATGGCAAGACCTCCAGGTGAATTTGAAGCTCACACAGATGTTA 875
Db 2245 AGAGCTGATGAACAAATGGCAAGACCTCCAGGTGAATTTGAAGCTCACACAGATGTTA 2304
QY 876 TCACAACCTGATGMAAACAAGCCAAAATCTGAGATCCCTGGAAGTTCCGATGATGC 935
Db 2305 TCACAACCTGATGMAAACAAGCCAAAATCTGAGATCCCTGGAAGTTCCGATGATGC 2364
QY 936 AGTCTGTTACAAAGAGCTTTGATTAACATGAACCTCAAGTGAAGTGAACCTCGAATAA 995
Db 2365 AGTCTGTTACAAAGAGCTTTGATTAACATGAACCTCAAGTGAAGTGAACCTCGAATAA 2424
QY 996 GTCTCTCAACATTAGTCCCATTTTGAAGCCAGTTCTGACAGTGAAGCCTCTGCACCT 1055
Db 2425 GTCTCTCAACATTAGTCCCATTTTGAAGCCAGTTCTGACAGTGAAGCCTCTGCACCT 2484
QY 1056 TTCTCTCAGAGAACTTCTGCTGTGCTACAGCTGAAGAAAGATGATGAATTAAGCCGACAGC 1115
Db 2485 TTCTCTCAGAGAACTTCTGCTGTGCTACAGCTGAAGAAAGATGATGAATTAAGCCGACAGC 2544
QY 1116 ACCTATTGAGAGCGACTTCCAGCAGTTCAAGACAGAAAGATGTACATAGAGCCTTCAA 1175
Db 2545 ACCTATTGAGAGCGACTTCCAGCAGTTCAAGACAGAAAGATGTACATAGAGCCTTCAA 2604
QY 1176 GAGGAATTGMAAATTAAGAAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATT 1235
Db 2605 GAGGAATTGMAAATTAAGAAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATT 2664
QY 1236 TCTGACAGAGACGCTTTTGAAGAGACTAGAGAACTCTACAGAGAGCCGACAGAGCTGCC 1295
Db 2665 TCTGACAGAGACGCTTTTGAAGAGACTAGAGAACTCTACAGAGAGCCGACAGAGCTGCC 2724
QY 1296 TCCTGAGGAGAGAGCCCAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGTCAA 1355
Db 2725 TCCTGAGGAGAGAGCCCAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGTCAA 2784
QY 1356 TACTGAGTGGAAAAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATGATGAGAC 1415
Db 2785 TACTGAGTGGAAAAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATGATGAGAC 2844
QY 1416 CTTGAAAGACTCCAGGAACCTTCAAGAGGCCAGATGAGCTGACCTCAAGCTGCGCA 1475
Db 2845 CTTGAAAGACTCCAGGAACCTTCAAGAGGCCAGATGAGCTGACCTCAAGCTGCGCA 2904
QY 1476 AGCTGAGGTGATCAAGGATCTGCGACGCCGTGGGCGATCTCTCATTTGACTCTCCA 1535
Db 2905 AGCTGAGGTGATCAAGGATCTGCGACGCCGTGGGCGATCTCTCATTTGACTCTCCA 2964
QY 1536 AGATCACTCGAGAAAGTCAAGGCACTTCAAGAGAAATTCGCTCTGAAGAGAACGT 1595
Db 2965 AGATCACTCGAGAAAGTCAAGGCACTTCAAGAGAAATTCGCTCTGAAGAGAACGT 3024
QY 1596 GAGCCAGTCAATGACCTTGTCTGCGCAGCTTACCACTTTGGGCAATTCAGCTCAACGTA 1655
Db 3025 GAGCCAGTCAATGACCTTGTCTGCGCAGCTTACCACTTTGGGCAATTCAGCTCAACGTA 3084
QY 1656 TAACCTCAGCACTCTGAGAGACCTGAACACAGATGAAAGCTTCTGCAAGTGGCCGTGCA 1715
Db 3085 TAACCTCAGCACTCTGAGAGACCTGAACACAGATGAAAGCTTCTGCAAGTGGCCGTGCA 3144
QY 1716 GAGCCAGTCAAGCAGCTGATGAAGCCACAGGACCTTGTCTCAGCATCTCAGCACTT 1775
Db 3145 GAGCCAGTCAAGCAGCTGATGAAGCCACAGGACCTTGTCTCAGCATCTCAGCACTT 3204
QY 1776 TCTTTCCAGCTGTCTCAAGGTTCCCTGGAGAGAGCCATCTCGCCAAAAGTGCCTTA 1835
Db 3205 TCTTTCCAGCTGTCTCAAGGTTCCCTGGAGAGAGCCATCTCGCCAAAAGTGCCTTA 3264
QY 1836 CTATATCAACACGAGACTCAAAACAACCTTGCTGGACCATCCCAAAATGACAGACTCTA 1895

Db 3265 CTATATCAACCAAGAGACTCAACAACACTTGTGGAGCCATCCCAAAATGACAGAGCTCTA 3324
QY 1896 CCAGTCTTTAGCTGACCTGAATATGTGAGATTTCTCAGCTTATAGACTGCGCATGAACCT 1955
Db 3325 CCAGTCTTTAGCTGACCTGAATATGTGAGATTTCTCAGCTTATAGACTGCGCATGAACCT 3384
QY 1956 CCGAAGACTGCAAGAGGCCCTTTGCTTGATCTCTTGAGCCTGTGAGCTGATGATGC 2015
Db 3385 CCGAAGACTGCAAGAGGCCCTTTGCTTGATCTCTTGAGCCTGTGAGCTGATGATGC 3444
QY 2016 CTTGGACCAAGCAACCTCAAGCAAAATGACCAAGCCCATGATATCCTGAGATTATTA 2075
Db 3445 CTTGGACCAAGCAACCTCAAGCAAAATGACCAAGCCCATGATATCCTGAGATTATTA 3504
QY 2076 TTGTTGACCACTATTATGACCGCTGAGCAAGACACAAATTTGTCACAGTCCC 2135
Db 3505 TTGTTGACCACTATTATGACCGCTGAGCAAGACACAAATTTGTCACAGTCCC 3564
QY 2136 TCTCTGCGTGATATGTGTGTAAGTGGCTGCTGATGTTTATGATACGGAGCAAGAG 2195
Db 3565 TCTCTGCGTGATATGTGTGTAAGTGGCTGCTGATGTTTATGATACGGAGCAAGAG 3624
QY 2196 GAGGATCCGTGCTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCATTTGGA 2255
Db 3625 GAGGATCCGTGCTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCATTTGGA 3684
QY 2256 AGACAAGTACAGATACCTTTTCAAGCAAGTGCAAGTTCACAGAGATTTGTGACCAAGC 2315
Db 3685 AGACAAGTACAGATACCTTTTCAAGCAAGTGCAAGTTCACAGAGATTTGTGACCAAGC 3744
QY 2316 CAGGCTGGGCTCTCTCTGATGATTTCTATCCAAATTCAGAGACAGTTGGGTGAAGTTC 2375
Db 3745 CAGGCTGGGCTCTCTCTGATGATTTCTATCCAAATTCAGAGACAGTTGGGTGAAGTTC 3804
QY 2376 ATCCTTTGGGGGAGTACATTGAGC 2401
Db 3805 ATCCTTTGGGGGAGTACATTGAGC 3830

RESULT 13

US-10-149-736-42
; Sequence 42, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149, 736
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238, 848
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 8689
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-42

Query Match 72.3%; Score 1735.6; DB 13; Length 8689;
Best local Similarity 98.6%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 628 TGGCAACATCTGTAGTAGACAGAGACCGCTGGTCTTTTACAGACACTCATAGA 687
Db 2948 TGGAGAAACATTCATAAAGGGTGAAGTGAAGAGAGAGGCTGCTTTGGAAGAACTCATAGA 3007

QY 688 TTACTGAACAGTTCCCCCTGACCTGGAAGATTCTTCTGCTGGCTTACAGAGCTGAA 747
Db 3008 TTACTGAACAGTTCCCCCTGACCTGGAAGATTCTTCTGCTGGCTTACAGAGCTGAA 3067
QY 748 ACACTGCCATGTCTCTACAGATGCTACCCCTAAGGAAGGCTCTAGAGACTCCAAG 807
Db 3068 ACACTGCCATGTCTCTACAGATGCTACCCCTAAGGAAGGCTCTAGAGACTCCAAG 3127
QY 808 GGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAGAGTGAATTGAAGCTCACACA 867
Db 3128 GGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAGAGTGAATTGAAGCTCACACA 3187
QY 868 GATGTTTATCAACACTTGATGAAGACCCAAATCTTGATATCCCTGGAAGTTC 927
Db 3188 GATGTTTATCAACACTTGATGAAGACCCAAATCTTGATATCCCTGGAAGTTC 3247
QY 928 GATGATGAGTCTCTTACAAAGACGTTTGATTAACATGAATCTCAAGTGAGTGAACCT 987
Db 3248 GATGATGAGTCTCTTACAAAGACGTTTGATTAACATGAATCTCAAGTGAGTGAACCT 3307
QY 988 CGAAGAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACAGTGAAGCGT 1047
Db 3308 CGAAGAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACAGTGAAGCGT 3367
QY 1048 CTGCACTTTCTCTGAGAACTTGTGTGCTACAGCTGAAGATGATGAATTAAAGC 1107
Db 3368 CTGCACTTTCTCTGAGAACTTGTGTGCTACAGCTGAAGATGATGAATTAAAGC 3427
QY 1108 CGGACAGCACTATTGAGGCGACTTTCCAGACAGTTCAAGACAGAAAGATGTACATAGG 1167
Db 3428 CGGACAGCACTATTGAGGCGACTTTCCAGACAGTTCAAGACAGAAAGATGTACATAGG 3487
QY 1168 GCCTTCAAGAGGGAATTGAATACTAAGAACTGTATCATGAGTACTCTTGAGACTGTA 1227
Db 3488 GCCTTCAAGAGGGAATTGAATACTAAGAACTGTATCATGAGTACTCTTGAGACTGTA 3547
QY 1228 CGAATATTTCTGACAGAGCAGCTTTGGAAGGACTAGAGAACTCTACAGAGAGCCCA 1287
Db 3548 CGAATATTTCTGACAGAGCAGCTTTGGAAGGACTAGAGAACTCTACAGAGAGCCCA 3607
QY 1288 GAGCTGCTCTCTGAGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 1347
Db 3608 GAGCTGCTCTCTGAGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 3667
QY 1348 GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCGGCTGACTGGCAGAGAAAAATA 1407
Db 3668 GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCGGCTGACTGGCAGAGAAAAATA 3727
QY 1408 GATGAGACCTTGAAGAGCTCCAGAACTTCAAGAGGCCAGGATGAGCTGAGCTCTCAAG 1467
Db 3728 GATGAGACCTTGAAGAGCTCCAGAACTTCAAGAGGCCAGGATGAGCTGAGCTCTCAAG 3787
QY 1468 CTGCGCAAGCTGAGGTGATCAAGGATCTCTGCGAGCCGCTGGCGATCTCTCATTTGAC 1527
Db 3788 CTGCGCAAGCTGAGGTGATCAAGGATCTCTGCGAGCCGCTGGCGATCTCTCATTTGAC 3847
QY 1528 TCTCTCAAGATCACTCGAAGAAAGTCAAGGCACTTCAAGAGGAAATTTGCGCTTGAAA 1587
Db 3848 TCTCTCAAGATCACTCGAAGAAAGTCAAGGCACTTCAAGAGGAAATTTGCGCTTGAAA 3907
QY 1588 GAGAACTGAGCCACGTCAATGACCTTGTCTGCGAGCTTACCACTTTGGCATTCAGCTC 1647
Db 3908 GAGAACTGAGCCACGTCAATGACCTTGTCTGCGAGCTTACCACTTTGGCATTCAGCTC 3967
QY 1648 TCACCGATAACCTCAGCACTCTGGAAGACCTGAACACAGATGGAAGCTTCTGAGGTG 1707
Db 3968 TCACCGATAACCTCAGCACTCTGGAAGACCTGAACACAGATGGAAGCTTCTGAGGTG 4027
QY 1708 GCCGTGAGGAGCCGAGTCAAGGAGCTGCATGAAGCCCAAGGAGCTTTGTCCAGACTCT 1767
Db 4028 GCCGTGAGGAGCCGAGTCAAGGAGCTGCATGAAGCCCAAGGAGCTTTGTCCAGACTCT 4087
QY 1768 CAGCACTTTCTTCCACGTGTCTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 1827

Db 4088 CAGCACTTCTTTCCACGTCTGTCCAGGGTCCCTGGAGAGAGCCATCTCGCCAAACAA 4147
QY 1828 GTGCCCTACTATATCAACACGAGACTCAACAACACTTGTCTGGACCATCCCAAAATGACA 1887
Db 4148 GTGCCCTACTATATCAACACGAGACTCAACAACACTTGTCTGGACCATCCCAAAATGACA 4207
QY 1888 GAGCTCTACCAAGTCTTTAGCTGACCTGTAATATGTCAAGTTCTCAGCTTATAGACTGCC 1947
Db 4208 GAGCTCTACCAAGTCTTTAGCTGACCTGTAATATGTCAAGTTCTCAGCTTATAGACTGCC 4267
QY 1948 ATGAACCTCCGAGAAGACTGCAGAAAGCCCTTTGCTTGGATCTCTTGAAGCCCTGTCAAGTCCA 2007
Db 4268 ATGAACCTCCGAGAAGACTGCAGAAAGCCCTTTGCTTGGATCTCTTGAAGCCCTGTCAAGTCCA 4327
QY 2008 TGTGATGCTTGGACCAAGCAACCACTCAAGCAAAATGACCAAGCCCATGGATATCTCTGAG 2067
Db 4328 TGTGATGCTTGGACCAAGCAACCACTCAAGCAAAATGACCAAGCCCATGGATATCTCTGAG 4387
QY 2068 ATTATTAATTGTTTGAACCACTATTATGACCCGCTGGAGCAAGACCAACAATTGGTTC 2127
Db 4388 ATTATTAATTGTTTGAACCACTATTATGACCCGCTGGAGCAAGACCAACAATTGGTTC 4447
QY 2128 AACGTCCCTCTCTGCGTGATATGTGTGAACCTGGCTGTGAATGTTATGATACGGGA 2187
Db 4448 AACGTCCCTCTCTGCGTGATATGTGTGAACCTGGCTGTGAATGTTATGATACGGGA 4507
QY 2188 CGAACAGGAGGATCCGTGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 2247
Db 4508 CGAACAGGAGGATCCGTGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 4567
QY 2248 CATTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGT 2307
Db 4568 CATTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGT 4627
QY 2308 GACCAGCGAGGCTGGGCTCTCTTCATGATTTCTATCCAAATTTCCAAAGACAGTTGGGT 2367
Db 4628 GACCAGCGAGGCTGGGCTCTCTTCATGATTTCTATCCAAATTTCCAAAGACAGTTGGGT 4687
QY 2368 GAAGTTGATCCTTTGGGGGAGTAACATTGAGC 2401
Db 4688 GAAGTTGATCCTTTGGGGGAGTAACATTGAGC 4721

RESULT 14
US-09-845-416-1
; Sequence 1, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11058
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-1

Query Match 72.3%; Score 1735.6; DB 13; Length 11058;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 628 TGGCAACATCTGTAGATGACAGAGACCGCTGGTCTTTTACAAGACACTCATAGA 687
Db 8008 TGGAGAGCATTCATAAAGGGTGAAGTGAAGAGGAGGCTGCTTTGGAAGAAACTCATAGA 8067

QY 688 TTAAGCAAGTTCCTCCCTGGACCTGGAAAAAGTTCTTGCTGGCTTACAGAGCTGAA 747
Db 8068 TTAAGCAAGTTCCTCCCTGGACCTGGAAAAAGTTCTTGCTGGCTTACAGAGCTGAA 8127
QY 748 ACAAGTCCAAATGCTCTACAGAGTGTACCCGTAAAGAAAGCTCTAGAGACTCCAAG 807
Db 8128 ACAAGTCCAAATGCTCTACAGAGTGTACCCGTAAAGAAAGCTCTAGAGACTCCAAG 8187
QY 808 GGAGTAAAGAGCTGATGAACAAATGGCAAGACCTCCAAAGTGAATTTGAAGCTCACACA 867
Db 8188 GGAGTAAAGAGCTGATGAACAAATGGCAAGACCTCCAAAGTGAATTTGAAGCTCACACA 8247
QY 868 GATGTTATCAACAACCTGATGAACAAAGCCAAAAATCTGAGATCCCTGGAAGTTCC 927
Db 8248 GATGTTATCAACAACCTGATGAACAAAGCCAAAAATCTGAGATCCCTGGAAGTTCC 8307
QY 928 GATGATGAGTCTCTGTACAAAGAGTTGGATTAACATGAATTTCAAGTGAAGTGAATCTT 987
Db 8308 GATGATGAGTCTCTGTACAAAGAGTTGGATTAACATGAATTTCAAGTGAAGTGAATCTT 8367
QY 988 CGAAAAAGTCTCTCAACATTAAGTCCCATTTGGAAGCCAGTTCTGACAGTGAAGGCT 1047
Db 8368 CGAAAAAGTCTCTCAACATTAAGTCCCATTTGGAAGCCAGTTCTGACAGTGAAGGCT 8427
QY 1048 CTGACCTTTCTCTGACAGAACTCTGCTGTGCTACAGCTGAAGATGATGAATTAAGC 1107
Db 8428 CTGACCTTTCTCTGACAGAACTCTGCTGTGCTACAGCTGAAGATGATGAATTAAGC 8487
QY 1108 CGCAGGACCTATTGGAGGCGACTTTCCAGCAGTTCAAGACAGATGTACATAGG 1167
Db 8488 CGCAGGACCTATTGGAGGCGACTTTCCAGCAGTTCAAGACAGATGTACATAGG 8547
QY 1168 GCCTTCAGAGGGAATTGAAACTTAAAGAACTGTATCATGAGTACTCTTGAGACTGTA 1227
Db 8548 GCCTTCAGAGGGAATTGAAACTTAAAGAACTGTATCATGAGTACTCTTGAGACTGTA 8607
QY 1228 CGAATATTTCTGACAGAGAGGCTTTGGAAGACTAGAAACTCTACAGAGAGCCGACA 1287
Db 8608 CGAATATTTCTGACAGAGAGGCTTTGGAAGACTAGAAACTCTACAGAGAGCCGACA 8667
QY 1288 GAGTGCCTCTGAGAGAGAGAGCCGAGATGTCACTCGGCTTTCTACGAAAGCAGGCTGAG 1347
Db 8668 GAGTGCCTCTGAGAGAGAGAGCCGAGATGTCACTCGGCTTTCTACGAAAGCAGGCTGAG 8727
QY 1348 GAGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 1407
Db 8728 GAGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 8787
QY 1408 GATGAGACCTTGAAAGACTCCAGGAACCTTCAAGAGGCCAGATGAGCTGAGACTCAAG 1467
Db 8788 GATGAGACCTTGAAAGACTCCAGGAACCTTCAAGAGGCCAGATGAGCTGAGACTCAAG 8847
QY 1468 CTGCGCAAGCTGAGGTGATCAAGGATCTCTGGCAGCCCGTGGGCGATCTCTCATTTGAC 1527
Db 8848 CTGCGCAAGCTGAGGTGATCAAGGATCTCTGGCAGCCCGTGGGCGATCTCTCATTTGAC 8907
QY 1528 TCTCTCAGATCACTCGAAGAACTCAAGGACCTTCAAGAGAAATTTGCGCTCTGAAA 1587
Db 8908 TCTCTCAGATCACTCGAAGAACTCAAGGACCTTCAAGAGAAATTTGCGCTCTGAAA 8967
QY 1588 GAGAAGTGAAGCAGTCAATGACCTTCTGCGCAGCTTACCACTTTGGGCAATTGAGCTC 1647
Db 8968 GAGAAGTGAAGCAGTCAATGACCTTCTGCGCAGCTTACCACTTTGGGCAATTGAGCTC 9027
QY 1648 TCACCGTATACTCAGACCTCTGGAAGACCTGAACACAGATGGAAGCTTCTGAGGTG 1707
Db 9028 TCACCGTATACTCAGACCTCTGGAAGACCTGAACACAGATGGAAGCTTCTGAGGTG 9087
QY 1708 GCCGTGAGAGCCGAGTCAAGGAGCTGATGAAGCCCAAGGAGCTTTGGTCCAGCATCT 1767
Db 9088 GCCGTGAGAGCCGAGTCAAGGAGCTGATGAAGCCCAAGGAGCTTTGGTCCAGCATCT 9147
QY 1768 CAGCACTTCTTTCCACGCTGTCTCAGAGGTCCCTGGAGAGAGCCATCTCGCCAAACAAA 1827

Db 9148 CAGCACTTTCTTCCACGCTGTGTCCAGGGTCCCTGGAGAGAGCCATCTCGCCAAACAA 9207
QY 1828 GTGCCCTACTATATCAACCAAGAGACTCAACCAACTTGTGGAGCCATCCCAAAATGACA 1887
Db 9208 GTGCCCTACTATATCAACCAAGAGACTCAACCAACTTGTGGAGCCATCCCAAAATGACA 9267
QY 1888 GAGCTCTAACAGTCTTTAGCTGACCTGAATATGTGAGATTTCTAGCTTATAGACTGCC 1947
Db 9268 GAGCTCTAACAGTCTTTAGCTGACCTGAATATGTGAGATTTCTAGCTTATAGACTGCC 9327
QY 1948 ATGAACTCCGAAGACTGCAAGAGGCCCTTGTGATCTCTGAGCCTGTCACTGCA 2007
Db 9328 ATGAACTCCGAAGACTGCAAGAGGCCCTTGTGATCTCTGAGCCTGTCACTGCA 9387
QY 2008 TGTGATGCTTGGACCAAGCAACCTCAAGCAAAATGACCAAGCCCATGATATCTGAG 2067
Db 9388 TGTGATGCTTGGACCAAGCAACCTCAAGCAAAATGACCAAGCCCATGATATCTGAG 9447
QY 2068 ATTATTAATTGTTGACCACTATTATGACCGCCTGAGCAAGACACAAATTTGGTC 2127
Db 9448 ATTATTAATTGTTGACCACTATTATGACCGCCTGAGCAAGACACAAATTTGGTC 9507
QY 2128 AACGTCCCTCTCTGCGTGATATGTGTGAACTGGCTGCTGAATGTTATGATACGGGA 2187
Db 9508 AACGTCCCTCTCTGCGTGATATGTGTGAACTGGCTGCTGAATGTTATGATACGGGA 9567
QY 2188 CGAACAGGAGGATCCGCTGTCTCTTTTAAACTGCGCATCATTTCCCTGTGTAAAGCA 2247
Db 9568 CGAACAGGAGGATCCGCTGTCTCTTTTAAACTGCGCATCATTTCCCTGTGTAAAGCA 9627
QY 2248 CATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGAGATTTGT 2307
Db 9628 CATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGAGATTTGT 9687
QY 2308 GACCAGCGCAGGCTGGGCTCTCTTCTGATGATTTCTATCCAAATTTCAAGACAGTTGGGT 2367
Db 9688 GACCAGCGCAGGCTGGGCTCTCTTCTGATGATTTCTATCCAAATTTCAAGACAGTTGGGT 9747
QY 2368 GAAGTTGATCCTTTGGGGCGAGTAACATTGAGC 2401
Db 9748 GAAGTTGATCCTTTGGGGCGAGTAACATTGAGC 9781

RESULT 15

US-10-149-736-44

; Sequence 44, Application US/10149736

; Publication No. US20030216332A1

; GENERAL INFORMATION:

; APPLICANT: Chamberlain, Jeffrey S.

; APPLICANT: Harper, Scott Q.

; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences

; FILE REFERENCE: UM-06968

; CURRENT APPLICATION NUMBER: US/10/149, 736

; CURRENT FILING DATE: 2002-06-17

; PRIOR APPLICATION NUMBER: PCT/US01/31126

; PRIOR FILING DATE: 2001-10-04

; PRIOR APPLICATION NUMBER: 60/238, 848

; PRIOR FILING DATE: 2000-10-06

; NUMBER OF SEQ ID NOS: 96

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 44

; LENGTH: 11443

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic

US-10-149-736-44

Query Match 72.3%; Score 1735.6; DB 13; Length 11443;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 628 TGGCAACACTCTGTAGATGACAGAGACCCTGGGTTCTTTTACAGACACTCATAGA 687
Db 5702 TGGAGAGAGCATTCATTAAGGGGTGAGTGAAGAGAGGCTGCTTTGGAGAAACTCATAGA 5761
QY 688 TTACTGCACAGTTCCCTTGGACTGGAAAAAGTTCTTGGCTGGCTTACAGAGCTGAA 747
Db 5762 TTACTGCACAGTTCCCTTGGACTGGAAAAAGTTCTTGGCTGGCTTACAGAGCTGAA 5821
QY 748 ACACTGCCAATGTCTTACAGAGATGTAACCCGTAAAGAAAGGCTCTAGAAAGACTCCAA 807
Db 5822 ACACTGCCAATGTCTTACAGAGATGTAACCCGTAAAGAAAGGCTCTAGAAAGACTCCAA 5881
QY 808 GGAGTAAAGAGCTGATGAACCAATGGCAAGACCTCCAAAGTGAATTGAAGCTCACACA 867
Db 5882 GGAGTAAAGAGCTGATGAACCAATGGCAAGACCTCCAAAGTGAATTGAAGCTCACACA 5941
QY 868 GATGTTATCACAACCTGATGAAAAAGCAAGCAAAAAATCTGAGATCCCTGGAAGTTCC 927
Db 5942 GATGTTATCACAACCTGATGAAAAAGCAAGCAAAAAATCTGAGATCCCTGGAAGTTCC 6001
QY 928 GATGATGAGTCTCTGTTACAAAGAGCTTTGATTAACATGAATCTCAAGTGAAGTAACTT 987
Db 6002 GATGATGAGTCTCTGTTACAAAGAGCTTTGATTAACATGAATCTCAAGTGAAGTAACTT 6061
QY 988 CGAATAAGTCTCTCAACATTAGTCCCATTTTGAAGCCAGTTCTGACCAAGTGAAGGCT 1047
Db 6062 CGAATAAGTCTCTCAACATTAGTCCCATTTTGAAGCCAGTTCTGACCAAGTGAAGGCT 6121
QY 1048 CTGACACTTCTCTGACAGAACTTCTGCTGTGCTACAGCTGAAGATGATGAATTAAC 1107
Db 6122 CTGACACTTCTCTGACAGAACTTCTGCTGTGCTACAGCTGAAGATGATGAATTAAC 6181
QY 1108 CGGCAAGCACTATTGAGAGGCACTTTCCAGCAGTTGCAAGACAGACAGATGTACATAGG 1167
Db 6182 CGGCAAGCACTATTGAGAGGCACTTTCCAGCAGTTGCAAGACAGACAGATGTACATAGG 6241
QY 1168 GCCTTCAAGAGGAATTGAAACTAAAGAACCTGTATCATGATCTCTTGAAGCTGTA 1227
Db 6242 GCCTTCAAGAGGAATTGAAACTAAAGAACCTGTATCATGATCTCTTGAAGCTGTA 6301
QY 1228 CGAATATTTCTGACAGAGCAGCTTTGGAAGGACTAGAGAACTCTACAGAGAGCCAGA 1287
Db 6302 CGAATATTTCTGACAGAGCAGCTTTGGAAGGACTAGAGAACTCTACAGAGAGCCAGA 6361
QY 1288 GAGTGCCTCTCTGAGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAGAGGCTGAG 1347
Db 6362 GAGTGCCTCTCTGAGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAGAGGCTGAG 6421
QY 1348 GAGTCAATACTGATGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 1407
Db 6422 GAGTCAATACTGATGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 6481
QY 1408 GATGAGACCTTGAAGAGCTCCAGAACTTCAAGAGGCCACGATGAGCTGAGCTCAAG 1467
Db 6482 GATGAGACCTTGAAGAGCTCCAGAACTTCAAGAGGCCACGATGAGCTGAGCTCAAG 6541
QY 1468 CTGGCCCAAGCTGAGGTGATCAAGGATCTCTGGAGCCCGTGGCGATCTCTCATTTGAC 1527
Db 6542 CTGGCCCAAGCTGAGGTGATCAAGGATCTCTGGAGCCCGTGGCGATCTCTCATTTGAC 6601
QY 1528 TCTCTCCAAGATCACTCGAGAAAGTCAAGGCACTTGAAGAGAAATTTGGCCTGAAA 1587
Db 6602 TCTCTCCAAGATCACTCGAGAAAGTCAAGGCACTTGAAGAGAAATTTGGCCTGAAA 6661
QY 1588 GAGAACGTGAGCCACGTCAATGACCTTGTCTGCGACCTTACCACTTTGGGCAATTAGCTC 1647
Db 6662 GAGAACGTGAGCCACGTCAATGACCTTGTCTGCGACCTTACCACTTTGGGCAATTAGCTC 6721
QY 1648 TCAACGTATACTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTG 1707
Db 6722 TCAACGTATACTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTG 6781
QY 1708 GCCGTGAGGAGCCAGTCAAGGAGCTGATGAAGCCCAAGGAGCTTTGGTCCAGCATCT 1767


```
Db 6782 GCGGTGAGGAGCCGAGTCAAGGCACTGATGAAGCCCAAGGGAATTGGTCCAGCATCT 6841
QY 1768 CAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGAGAGAGCCATCTCGCCAAACAAA 1827
Db 6842 CAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGAGAGAGCCATCTCGCCAAACAAA 6901
QY 1828 GTGCCCTACTATATCAACCAAGAGACTCAAAACAATTGCTGGAGACCATCCCAAAATGACA 1887
Db 6902 GTGCCCTACTATATCAACCAAGAGACTCAAAACAATTGCTGGAGACCATCCCAAAATGACA 6961
QY 1888 GAGCTCTACCAAGTCTTAGCTGACCTGAATAATGTCAAGATTCTCAGCTTATAGGACTGCC 1947
Db 6962 GAGCTCTACCAAGTCTTAGCTGACCTGAATAATGTCAAGATTCTCAGCTTATAGGACTGCC 7021
QY 1948 ATGAAGCTCCGAGAAGCTGCAGAGGCTTTGCTTGATCTCTTGAGCCTGTCACTGCA 2007
Db 7022 ATGAAGCTCCGAGAAGCTGCAGAGGCTTTGCTTGATCTCTTGAGCCTGTCACTGCA 7081
QY 2008 TGTGATGCCCTTGAACCAAGCAAACTCAAGCAAAATGACCAAGCCATGATATCTGCAG 2067
Db 7082 TGTGATGCCCTTGAACCAAGCAAACTCAAGCAAAATGACCAAGCCATGATATCTGCAG 7141
QY 2068 ATTATTAATGTTTGACCACTATTTAAGACCGCTGAGCAGAGACACAACAATTGGTC 2127
Db 7142 ATTATTAATGTTTGACCACTATTTAAGACCGCTGAGCAGAGACACAACAATTGGTC 7201
QY 2128 AACGTCCCTCTCTGCGTGATATGTCTGAAGTGGCTGTAATGTTATGATACGGGA 2187
Db 7202 AACGTCCCTCTCTGCGTGATATGTCTGAAGTGGCTGTAATGTTATGATACGGGA 7261
QY 2188 CGAACAGGAGATCCGTGTCTCTTTTAAACTGGCATATTCCCTGTGTAAAGCA 2247
Db 7262 CGAACAGGAGATCCGTGTCTCTTTTAAACTGGCATATTCCCTGTGTAAAGCA 7321
QY 2248 CATTTGAGAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTGT 2307
Db 7322 CATTTGAGAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTGT 7381
QY 2308 GACCAGCGCAGGCTGGGCTCCTTCTGATGATTCTATCCAAATTCCAAGACAGTTGGGT 2367
Db 7382 GACCAGCGCAGGCTGGGCTCCTTCTGATGATTCTATCCAAATTCCAAGACAGTTGGGT 7441
QY 2368 GAAGTTGCATCCTTTGGGGGCAAGTAACATTGAGC 2401
Db 7442 GAAGTTGCATCCTTTGGGGGCAAGTAACATTGAGC 7475
```

Search completed: February 2, 2004, 17:38:52
Job time : 806.082 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 11:11:40 ; Search time 148.424 Seconds
(without alignments)
7140.092 Million cell updates/sec

Title: US-09-845-416-9_COPY_1000_3400

Perfect score: 2401

Sequence: 1 ggccagtcattgatgagagag.....tggggggcagtaacattgagc 2401

Scoring table: IDENTITY_NUC

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|---------|-------|---------------------|
| 1 | 1735.6 | 72.3 | 5952 | 4 | US-09-687-875A-1 |
| 2 | 1724.6 | 71.8 | 13977 | 4 | US-09-484-970B-60 |
| 3 | 1506.8 | 62.8 | 19307 | 3 | US-08-836-022A-10 |
| 4 | 1506.8 | 62.8 | 19307 | 3 | US-09-427-048A-10 |
| 5 | 690.6 | 28.8 | 6045 | 4 | US-09-091-501B-7 |
| 6 | 690.6 | 28.8 | 10320 | 4 | US-09-091-501B-9 |
| 7 | 79.4 | 3.3 | 200 | 4 | US-09-091-501B-5 |
| 8 | 78.6 | 3.3 | 200 | 4 | US-09-091-501B-4 |
| 9 | 78.6 | 3.3 | 200 | 4 | US-09-091-501B-6 |
| 10 | 76.6 | 3.2 | 7218 | 1 | US-08-232-463-14 |
| 11 | 63.6 | 2.6 | 238 | 4 | US-09-687-875A-13 |
| 12 | 44.2 | 1.8 | 2574 | 4 | US-09-668-313A-10 |
| 13 | 44 | 1.8 | 1230025 | 4 | US-09-198-452A-1 |
| 14 | 43.4 | 1.8 | 1179 | 4 | US-09-107-532A-1186 |
| 15 | 42.8 | 1.8 | 1690 | 4 | US-09-620-312D-69 |
| 16 | 42.8 | 1.8 | 7812 | 3 | US-09-368-590-1 |
| 17 | 40.4 | 1.7 | 2223 | 1 | US-08-257-073-4 |
| 18 | 39.2 | 1.6 | 16995 | 4 | US-08-961-527-82 |
| 19 | 38.6 | 1.6 | 1751 | 4 | US-09-620-312D-847 |
| 20 | 38.6 | 1.6 | 1995 | 1 | US-08-425-069-3 |
| 21 | 38.6 | 1.6 | 1995 | 2 | US-08-317-844B-3 |
| 22 | 38.4 | 1.6 | 7672 | 4 | US-09-220-132-24 |
| 23 | 38.2 | 1.6 | 428 | 4 | US-09-668-313A-3 |
| 24 | 38.2 | 1.6 | 1131 | 6 | 5180810-3 |
| 25 | 38.2 | 1.6 | 1784 | 6 | 5180810-2 |
| 26 | 38.2 | 1.6 | 4439 | 4 | US-09-668-313A-17 |
| 27 | 38 | 1.6 | 1394 | 4 | US-09-247-155-76 |

| | | | | | | |
|----|------|-----|--------|---|--------------------|--------------------|
| 28 | 36.8 | 1.5 | 1886 | 6 | 5210183-1 | Patent No. 5210183 |
| 29 | 36.6 | 1.5 | 1845 | 4 | US-08-887-534A-22 | Sequence 22, Appl |
| 30 | 36.6 | 1.5 | 1845 | 4 | US-09-527-431-22 | Sequence 22, Appl |
| 31 | 36.2 | 1.5 | 289 | 3 | US-09-007-005-17 | Sequence 17, Appl |
| 32 | 36.2 | 1.5 | 289 | 3 | US-09-244-796-17 | Sequence 17, Appl |
| 33 | 36.2 | 1.5 | 2447 | 2 | US-09-014-969-14 | Sequence 14, Appl |
| 34 | 36.2 | 1.5 | 168575 | 4 | US-09-426-290-1 | Sequence 1, Appli |
| 35 | 36 | 1.5 | 2873 | 4 | US-08-630-915A-193 | Sequence 193, App |
| 36 | 35.8 | 1.5 | 790 | 3 | US-09-461-474-13 | Sequence 13, Appl |
| 37 | 35.6 | 1.5 | 4868 | 1 | US-08-139-937-12 | Sequence 12, Appl |
| 38 | 35.6 | 1.5 | 4868 | 5 | PCT-US93-11310-12 | Sequence 12, Appl |
| 39 | 35.6 | 1.5 | 8257 | 4 | US-09-595-684B-30 | Sequence 30, Appl |
| 40 | 35.6 | 1.5 | 8789 | 1 | US-08-328-254-5 | Sequence 5, Appli |
| 41 | 35.6 | 1.5 | 10136 | 1 | US-08-353-700-2 | Sequence 2, Appli |
| 42 | 35.6 | 1.5 | 10136 | 5 | PCT-US95-16216-2 | Sequence 2, Appli |
| 43 | 35.2 | 1.5 | 3466 | 1 | US-08-468-036-38 | Sequence 38, Appl |
| 44 | 35.2 | 1.5 | 3466 | 2 | US-08-376-843-38 | Sequence 38, Appl |
| 45 | 35 | 1.5 | 2169 | 4 | US-09-434-408-3 | Sequence 3, Appli |

ALIGNMENTS

RESULT 1
US-09-687-875A-1
; Sequence 1, Application US/09687875A
; Patent No. 6544786
; GENERAL INFORMATION:
; APPLICANT: Xieo, Xieo
; APPLICANT: Liu, Paul
; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPICEI
; FILE REFERENCE: 00792
; CURRENT APPLICATION NUMBER: US/09/687, 875A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/158, 868
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 5952
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2897)..(2898)
; OTHER INFORMATION: S4 junction site
; NAME/KEY: misc feature
; LOCATION: (3198)..(3199)
; OTHER INFORMATION: S2 junction site
US-09-687-875A-1

Query Match 72.3%; Score 1735.6; DB 4; Length 5952;
Best local Similarity 98.6%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

| | | | |
|----|------|---|------|
| QY | 628 | TGGGCAACATCTGTAGATGACAGAGACCGCTGGTCTTTTACAGACACTCATAGA | 687 |
| DB | 2902 | TGAGAGCATTCATAAAGGCTGAGTGAAGCGAGGCTGCTTTGGAAGAACTCATAGA | 2961 |
| QY | 688 | TTACTGCAACAGTTCCTCCCTGAGCCTGMAAAGTTCTTGCTGCTTACAGAACTGA | 747 |
| DB | 2962 | TTACTGCAACAGTTCCTCCCTGAGCCTGMAAAGTTCTTGCTGCTTACAGAACTGA | 3021 |
| QY | 748 | ACAACTGCCAATGTCTTACAGAGTCTACCCGTAAAGAAAGGCTCTAGAACTCCAAG | 807 |
| DB | 3022 | ACAACTGCCAATGTCTTACAGAGTCTACCCGTAAAGAAAGGCTCTAGAACTCCAAG | 3081 |
| QY | 808 | GGAGTAAAGAGCTGTATGAACAATGGCAAGACCTCCAAGGTGAATTGAAGCTCACAA | 867 |
| DB | 3082 | GGAGTAAAGAGCTGTATGAACAATGGCAAGACCTCCAAGGTGAATTGAAGCTCACAA | 3141 |
| QY | 868 | GATGTTATCAACACCTGATGAAGAAACAGCCAAAATCTGAGATCCCTGAAGTTCC | 927 |

```
Db 3142 GATGTTTATCACACCTGGATGAAAAACAGCCAAAAATCCTGAGATCCCTGGAAGTCC 3201
QY 928 GATGATGCAGTCTCTGTTACAAAGACGTTTGATATACATACTTCAAGTGAAGTGA 987
Db 3202 GATGATGCAGTCTCTGTTACAAAGACGTTTGATATACATACTTCAAGTGAAGTGA 3261
QY 988 CGGAAAAAGTCTCTCAACATTAAGTCCCATTTGGAGCCAGTTCTGACCAGTGAAGCGT 1047
Db 3262 CGGAAAAAGTCTCTCAACATTAAGTCCCATTTGGAGCCAGTTCTGACCAGTGAAGCGT 3321
QY 1048 CTGCACCTTCTCTGACGAACTTCTGTTGCTGCTACAGCTGAAGATGATGAATTAA 1107
Db 3322 CTGCACCTTCTCTGACGAACTTCTGTTGCTGCTACAGCTGAAGATGATGAATTAA 3381
QY 1108 CGGACGGACCTATTGGAGGCGACTTCCAGCAGTTCAAGACAGAACGATGTACATAG 1167
Db 3382 CGGACGGACCTATTGGAGGCGACTTCCAGCAGTTCAAGACGATGTACATAG 3441
QY 1168 GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAAGCTGTA 1227
Db 3442 GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAAGCTGTA 3501
QY 1228 CGAATATTTCTGACAGACAGCAGCTTTGGAAGACCTAGAGAACTCTACCAAGAGCCAGA 1287
Db 3502 CGAATATTTCTGACAGACAGCAGCTTTGGAAGACCTAGAGAACTCTACCAAGAGCCAGA 3561
QY 1288 GAGCTGCTCTGAGAGAGAGAGCCAGATGTCACTGCGCTTCTACGAAAGCAGGCTGAG 1347
Db 3562 GAGCTGCTCTGAGAGAGAGAGCCAGATGTCACTGCGCTTCTACGAAAGCAGGCTGAG 3621
QY 1348 GAGTCAATATCTAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 1407
Db 3622 GAGTCAATATCTAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 3681
QY 1408 GATGAGACCCCTTGAAGAAGCTCAGAGAACTTCAAGAGCCACGATGAGCTGGAACCTCAAG 1467
Db 3682 GATGAGACCCCTTGAAGAAGCTCAGAGAACTTCAAGAGCCACGATGAGCTGGAACCTCAAG 3741
QY 1468 CTGCGCCAAGCTGAGGTGATCAAGGATCTCTGCGAGCCGCTGGCGATCTCTCATTTGAC 1527
Db 3742 CTGCGCCAAGCTGAGGTGATCAAGGATCTCTGCGAGCCGCTGGCGATCTCTCATTTGAC 3801
QY 1528 TCTCTCAAGATCACCTCGAGAAAGTCAAGGCACTTCAAGAGAGAAATTGCGCTCTGAAA 1587
Db 3802 TCTCTCAAGATCACCTCGAGAAAGTCAAGGCACTTCAAGAGAGAAATTGCGCTCTGAAA 3861
QY 1588 GAGAAGCTGAGCCACGTCATATGACCTTCTGCGCAGCTTAACTTTGGGCACTTCAGCTC 1647
Db 3862 GAGAAGCTGAGCCACGTCATATGACCTTCTGCGCAGCTTAACTTTGGGCACTTCAGCTC 3921
QY 1648 TCACCGTATTAACCTCAGCAGCTTGGAGAACCTGAACACAGATGGAAGCTTCTGCAAGTG 1707
Db 3922 TCACCGTATTAACCTCAGCAGCTTGGAGAACCTGAACACAGATGGAAGCTTCTGCAAGTG 3981
QY 1708 GCGGTGAGGAGCCGAGTCAGGAGCTGCATGAAGCCACAGGGACTTTGTCAGCATCT 1767
Db 3982 GCGGTGAGGAGCCGAGTCAGGAGCTGCATGAAGCCACAGGGACTTTGTCAGCATCT 4041
QY 1768 CAGCACTTCTTTCCACGCTCTGTCAGGGTCCCTGGAGAGAGCCATCTGCGCAAAACAAA 1827
Db 4042 CAGCACTTCTTTCCACGCTCTGTCAGGGTCCCTGGAGAGAGCCATCTGCGCAAAACAAA 4101
QY 1828 GTGCCCTACTATATCAACACAGAGACTCAAAACAACTTCTGCGGAGCCATCCCAAAATGACA 1887
Db 4102 GTGCCCTACTATATCAACACAGAGACTCAAAACAACTTCTGCGGAGCCATCCCAAAATGACA 4161
QY 1888 GAGCTCTACCAAGTCTTGAAGTCACTGAATATGTCAGATCTCAGCTTATAGAGACTGCC 1947
Db 4162 GAGCTCTACCAAGTCTTGAAGTCACTGAATATGTCAGATCTCAGCTTATAGAGACTGCC 4221
QY 1948 ATGAAGTCCGAAAGACTGCAGAGGCCCTTTGCTTGATCTCTTGAGCTGTCAAGTGA 2007
Db 4222 ATGAAGTCCGAAAGACTGCAGAGGCCCTTTGCTTGATCTCTTGAGCTGTCAAGTGA 4281
```

```
QY 2008 TGTGATGCTTGGACCCAGACCAACCTCAAGCAAAATGACAGCCCATGATATCTGACG 2067
Db 4282 TGTGATGCTTGGACCCAGACCAACCTCAAGCAAAATGACAGCCCATGATATCTGACG 4341
QY 2068 ATTATTAATTGTTGACCACTATTATATGACCGCTGGAGCAAGAGACACAATTGGTC 2127
Db 4342 ATTATTAATTGTTGACCACTATTATATGACCGCTGGAGCAAGAGACACAATTGGTC 4401
QY 2128 AACGTCCTCTCTGCGTGAATATGTCGTAAGTGGCTGTAATGTTATGATACGGGA 2187
Db 4402 AACGTCCTCTCTGCGTGAATATGTCGTAAGTGGCTGTAATGTTATGATACGGGA 4461
QY 2188 CGAACAGGAGATCCGTCCTCTGCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCA 2247
Db 4462 CGAACAGGAGATCCGTCCTCTGCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCA 4521
QY 2248 CATTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAAGAGATTGGT 2307
Db 4522 CATTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAAGAGATTGGT 4581
QY 2308 GACCAAGCAGAGCTGGGCTCTCTGATGATCTATCCAAATTCCAAAGAGTTGGGT 2367
Db 4582 GACCAAGCAGAGCTGGGCTCTCTGATGATCTATCCAAATTCCAAAGAGAGTTGGGT 4641
QY 2368 GAAGTTGATCTCTTTGGGGCAGTAAATTTGAGC 2401
Db 4642 GAAGTTGATCTCTTTGGGGCAGTAAATTTGAGC 4675
```

RESULT 2

```
US-09-484-970B-60
; Sequence 60, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmuth, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELLING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484, 970B
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 60
; LENGTH: 13977
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6426186 229357.11CBI
; LOCATION: 11721-11761, 12294, 13969
; OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-60
```

```
Query Match 71.8%; Score 1724.6; DB 4; Length 13977;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 24; Indels 1; Gaps 1;
```

```
QY 628 TGGGCAACATCTGTATGAGACAGAGACCGCTGGGTTCTTTACAGACACTCATAGA 687
Db 8216 TGGAGAAACATTCATATAAAGGGTAGTGAAGCGAGAGGCTGCTTTGGAAGAACTCATAGA 8275
QY 688 TTAAGTCAACAGTTCCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 747
Db 8276 TTAAGTCAACAGTTCCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 8335
QY 748 ACAAGTCCCAATGTCTTACAGAGATGCTACCCGTAAGAGAAAGGCTCTAGAGACTCCAAG 807
Db 8336 ACAAGTCCCAATGTCTTACAGAGATGCTACCCGTAAGAGAAAGGCTCTAGAGACTCCAAG 8395
QY 808 GGAATAAAGAGCTGATGAACAATGCAAGAGCTCCAAAGGTGAATTTGAAGCTCACACA 867
```


Db 8396 GGAGTAAAGAGCTGATGAACATGGCAAGACCTCCAGGTGAATTTGAAGCTCACACA 8455
Qy 868 GATGTTATCAACAACCTGATGAAAAACAGCCAAAAATCCTGAGATCCCTGGAAGTTCC 927
Db 8456 GATGTTATCAACAACCTGATGAAAAACAGCCAAAAATCCTGAGATCCCTGGAAGTTCC 8515
Qy 928 GATGATGAGTCTCTGTTACAAAGACGTTTGATTAACATGAATTTCAAGTGAAGTGA 987
Db 8516 GATGATGAGTCTCTGTTACAAAGACGTTTGATTAACATGAATTTCAAGTGAAGTGA 8575
Qy 988 CGAATAAGTCTCTCAACATTAAGTCCCATTTTGAAGCCAGTTCTGACCAGTGAAGCCT 1047
Db 8576 CGAATAAGTCTCTCAACATTAAGTCCCATTTTGAAGCCAGTTCTGACCAGTGAAGCCT 8635
Qy 1048 CTGCACCTTTCTCTGACAGAACTCTGCTGCTGCTACAGCTGAAGATGATGAATTAAG 1107
Db 8636 CTGCACCTTTCTCTGACAGAACTCTGCTGCTGCTACAGCTGAAGATGATGAATTAAG 8695
Qy 1108 CGGCAAGCACTTATTTGAGGCGACTTTCCAGCAGTTCCAGAACGAAACGATGATAGG 1167
Db 8696 CGGCAAGCACTTATTTGAGGCGACTTTCCAGCAGTTCCAGAACGAAACGATGATAGG 8755
Qy 1168 GCCTTCAAGAGGGAATTTGAAACTAAAGAACTGTATCATGAGTACTTTGAGACTGTA 1227
Db 8756 GCCTTCAAGAGGGAATTTGAAACTAAAGAACTGTATCATGAGTACTTTGAGACTGTA 8815
Qy 1228 CGAATATTTTGAACAGAGCAGCTTTGGAAGGACTGAGAACTCTACAGAGAGCCAGA 1287
Db 8816 CGAATATTTTGAACAGAGCAGCTTTGGAAGGACTGAGAACTCTACAGAGAGCCAGA 8875
Qy 1288 GAGTGCCTCTGAGAGAGAGAGCCAGAACTGCTGCTGCTTCTACGAAGCAGGCTGAG 1347
Db 8876 GAGTGCCTCTGAGAGAGAGAGCCAGAACTGCTGCTGCTTCTACGAAGCAGGCTGAG 8935
Qy 1348 GAGTCAATACTGAGTGGAAAAATTTGAACCTGCACCTCCGCTGACTGCGAGAAAAATA 1407
Db 8936 GAGTCAATACTGAGTGGAAAAATTTGAACCTGCACCTCCGCTGACTGCGAGAAAAATA 8995
Qy 1408 GATGAGACCTTTGAAGACTCCAGAACTTCAAGAGGCCAGATGAGCTGAGCTCAAG 1467
Db 8996 GATGAGACCTTTGAAGACTCCAGAACTTCAAGAGGCCAGATGAGCTGAGCTCAAG 9055
Qy 1468 CTGCGCAAGCTGAGGTGATCAAGGGATCCTGCGAGCCCTGGGGGATCTCTCATTTGAC 1527
Db 9056 CTGCGCAAGCTGAGGTGATCAAGGGATCCTGCGAGCCCTGGGGGATCTCTCATTTGAC 9115
Qy 1528 TCTCTCAAGATCACTCGAAGAACTCAAGGCACTTCAAGAGAAATTTGCCCTCTGAAA 1587
Db 9116 TCTCTCAAGATCACTCGAAGAACTCAAGGCACTTCAAGAGAAATTTGCCCTCTGAAA 9175
Qy 1588 GAGAACGTGAGCCACGTCATGACCTTGCTGCGCAGCTTACCACTTTGGGCATTCAGCTC 1647
Db 9176 GAGAACGTGAGCCACGTCATGACCTTGCTGCGCAGCTTACCACTTTGGGCATTCAGCTC 9235
Qy 1648 TCACCGTATTAACCTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGAGG 1707
Db 9236 TCACCGTATTAACCTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGAGG 9295
Qy 1708 GCCGTGAGGAGCCGAGTCAGGAGCTGATGAAGCCACAGGGAATTTGTTCCAGCATCT 1767
Db 9296 GCCGTGAGGAGCCGAGTCAGGAGCTGATGAAGCCACAGGGAATTTGTTCCAGCATCT 9355
Qy 1768 CAGCACTTTCTTCCACGCTGTCAGAGGCTCCCTGGAGAGAGCCATCTCGCCAAACAA 1827
Db 9356 CAGCACTTTCTTCCACGCTGTCAGAGGCTCCCTGGAGAGAGCCATCTCGCCAAACAA 9415
Qy 1828 GTGCCCTACTATATCAACACGAGACTCAACCACTTGTGGGACCATCCCAAAATGACA 1887
Db 9416 GTGCCCTACTATATCAACACGAGACTCAACCACTTGTGGGACCATCCCAAAATGACA 9475
Qy 1888 GAGCTTAACAGTCTTTAGCTGAGCTGAATATGTCAAGATTTCAAGCTTATAGAGCTGCC 1947

Db 9476 GAGCTCTAACAGTCTTTAGCTGACCTGAATATGTCAGATTTCTCAGCTTATAGAGCTGCC 9535
Qy 1948 ATGAACCTCCGAAGACTGACAGAGAGCCCTTTGCTTGGATCTCTTGAGCCCTGTACGTCGA 2007
Db 9536 ATGAACCTCCGAAGACTGACAGAGAGCCCTTTGCTTGGATCTCTTGAGCCCTGTACGTCGA 9595
Qy 2008 TGTGATGCTTGGACACAGCAACCTCAAGCAAAATGACAGCCCATGATATCTGACAG 2067
Db 9596 TGTGATGCTTGGACACAGCAACCTCAAGCAAAATGACAGCCCATGATATCTGACAG 9655
Qy 2068 ATTATTAATTTGTTGACCACTAATTATGACCGCCTGAGCAAGACCAACAATTTGCTC 2127
Db 9656 ATTATTAATTTGTTGACCACTAATTATGACCGCCTGAGCAAGACCAACAATTTGCTC 9715
Qy 2128 AACGTCCTCTCTGCGTGATATGTCCTGAACCTGCTGCTGAATGTTATGATACGGA 2187
Db 9716 AACGTCCTCTCTGCGTGATATGTCCTGAACCTGCTGCTGAATGTTATGATACGGA 9775
Qy 2188 CGAACAGGAGATCCGTCCTCTCTTTTAAACTGCGATCATTTCCCTGTGTAAGCA 2247
Db 9776 CGAACAGGAGATCCGTCCTCTCTTTTAAACTGCGATCATTTCCCTGTGTAAGCA 9835
Qy 2248 CATTTGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGAGATTTGT 2307
Db 9836 CATTTGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGAGATTTGT 9895
Qy 2308 GACCAGCGAGGCT-GGGCTCTCTCTGATGATTTCTTCCAATTTCCAAGACAGTTGGG 2366
Db 9896 GACCAGCGAGGCTGGGGCTCTCTTCTGATGATTTCTTCCAATTTCCAAGACAGTTGGG 9955
Qy 2367 TGAAGTTCATCTTTGGGGGCAATAATTGAGC 2401
Db 9956 TGAAGTTCATCTTTGGGGGCAATAATTGAGC 9990

RESULT 3
US-08-836-022A-10/c
; Sequence 10, Application US/08836022A
; Patent No. 6001557
; GENERAL INFORMATION:
; APPLICANT: Trustees of the University of Pennsylvania
; APPLICANT: Wilson, James M.
; APPLICANT: Fisher, Krishna J.
; APPLICANT: Chen, Shu-Jen
; APPLICANT: Weltzman, Matthew
; TITLE OF INVENTION: Improved Adenovirus Virus and
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P O Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,022A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/331,381
; FILING DATE: 28-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GNVN.008PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818

; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19307 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; US-08-836-022A-10

Query Match 62.8%; Score 1506.8; DB 3; Length 19307;
Best Local Similarity 90.6%; Pred. No. 0;
Matches 1607; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 628 TGGGCAACATCTGTAGATGACAGAGACCGCTGGGTTCTTTTACAGACACTCATAGA 687
Db 6478 TGGGGAACATTCATAAAGAGTAAGTACAGCAAGAGGCTGCTTTGGAAGAACTCATAGA 6419
QY 688 TTAAGTCAACAGTTCCTCCCTGGACCTGGAAGAGTTTCTGCTGCTTACAGAGCTGAA 747
Db 6418 TTAAGTCAACAGTTCCTCCCTGGACCTGGAAGAGTTTCTGCTGCTTACAGAGCTGAA 6359
QY 748 ACAAGTCCCAATGTCTCTACAGATGCTACCCGTAAGAGAAAGGCTCCTAGAGACTCCAAG 807
Db 6358 ACAAGTCCCAATGTCTCTACAGATGCTACCCGTAAGAGAAAGGCTCCTAGAGACTCCAAG 6299
QY 808 GGAGTAAAGAGCTGATGAACAATGGCAAGAGCTCCAGAGTGAATGAAGCTCACACA 867
Db 6298 GGAGTCAAGAGCTGATGAACAATGGCAAGAGCTCCAGAGTGAATGAATGAATCAACACA 6239
QY 868 GATGTTATCAACAACCTGATGAAGAAACAGCCAAAGAAATCCTGAGATCCCTGGAAGTCC 927
Db 6238 GATATCTATCAACATCTTGATGAAGAAATGGCCAAAGAAATCCTGAGATCCCTGGAAGTCCG 6179
QY 928 GATATGCAAGTCTCTGTACAAAGACGTTTGATTAACATGAAGTCAAGTGAAGTGAAGT 987
Db 6178 GATGAAGCAACCCCTGTACAAAGACGTTTGATTAACATGAAGTCAAGTGAAGTGAAGT 6119
QY 988 CGGAAAAAGTCTCTCAACATTAAGTCCCATTTGGAAGCCAGTTCTGACCAAGTGAAGCGT 1047
Db 6118 CAGAAAAAGTCTCTCAACATTAAGTCCCATTTGGAAGCCAGTTCTGACCAAGTGAAGCGT 6059
QY 1048 CTGCACTTTTCTCTGCAAGAACTTCTGCTGCTACAGCTGAAGAGTGAATTAAGC 1107
Db 6058 TTGCACTTTTCTCTCTGCAAGAACTTCTGCTGCTACAGCTGAAGAGTGAATTAAGC 5999
QY 1108 CGGCAAGCACTTATTTGAGGCACTTTCCAGCAGTTTCAAGAGCAAGCAAGCAAGTGAAG 1167
Db 5998 CGTCAAGCACTTATTTGAGGCACTTTCCAGCAGTTTCAAGAGCAAGCAAGTGAAGTGAAG 5939
QY 1168 GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTA 1227
Db 5938 GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTA 5879
QY 1228 CGAATATTTTCTGACAGAGCAGCTTTGGAAGGACTAGAGAACTCTACAGAGAGCCAGAG 1287
Db 5878 AGAATATTTTCTGACAGAGCAGCTTTGGAAGGACTAGAGAACTCTACAGAGAGCCAGAG 5819
QY 1288 GAGTGCCTCTGAGAGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 1347
Db 5818 GAAGTGCCTCTGAGAGAGAGAGCTCAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAA 5759
QY 1348 GAGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 1407
Db 5758 GAGTCAACGCTGAATGGGACAAATTGAACCTGCGCTCAAGCTGATTTGGCAGAGAAAAATA 5699
QY 1408 GATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGATGAGCTGACCTCAAG 1467
Db 5698 GATGAGCTCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGATGAGCTGACCTCAAG 5639
QY 1468 CTGCGCCAAAGCTGAGGTGATCAAGGAGTCTGCGAGCCGCTGGCGCATCTCTCATTTGAC 1527
Db 5638 TTGCGCCAAAGCTGAGGTGATCAAGGAGTCTGCGAGCCAGTGGGCGATCTCTCATTTGAC 5579

QY 1528 TCTCTCCAGATCACTCGAGAAAGTCAAGGCACTTTCAGAGAGAAATTGCGCTCTGAAA 1587
Db 5578 TCTCTCCAGATCACTCGAGAAAGTCAAGGCACTTTCAGAGAGAAATTGCACTCTTAAA 5519
QY 1588 GAGACGTGAGCCACGTCAATGAACCTTGTCTGCGCAGCTTACCACCTTTGGGCACTCAGCTC 1647
Db 5518 GAGATGTCAATCGTGTATGAACCTTGCACATCAGCTGACCACTGGGCACTTCAAGCTC 5459
QY 1648 TCAACGTATAACCTCAGACACTCTGGAAGACCTGAACACAGATGGAAGCTTCTGACAGTG 1707
Db 5458 TCACTTATAACCTCAGACACTTTCGAAGATCTGAATACAGATGAGGCTTCTACAGTG 5399
QY 1708 GCGGTGAGGAGCCGAGTCAAGCAGCTGATGAAGGCCACAGGAGCTTTGTCAGCATCT 1767
Db 5398 GCTGTGAGGAGCCGAGTCAAGCAGCTGATGAAGGCCACAGGAGCTTTGTCAGCATCT 5339
QY 1768 CAGCACTTTCTTCCAGGCTGTCCAGAGTCCCTGGGAGAGAGCCATCTGCAAAACAAA 1827
Db 5338 CAGCACTTCTTCCAGTCAAGTCAAGGCTCCCTGGGAGAGAGCCATCTGCAAAACAAA 5279
QY 1828 GTGCCCTACTATATCAACAGAGACTCAAACTTGCTGGAGACCATCCCAAAATGACA 1887
Db 5278 GTGCCCTACTATATCAACAGAGACTCAAACTTGCTGGAGACCATCCCAAAATGACA 5219
QY 1888 GAGCTCTACAGTCTTTAGCTGACCTGAATATGTCAGATTTCTCAGCTTATAGACTGCC 1947
Db 5218 GAGCTCTACAGTCTTTAGCTGACCTGAATATGTCAGATTTCTCAGCTTATAGACTGCC 5159
QY 1948 ATGAAGCTCCGAGAGCTGAGAGAGGCCCTTTGCTTGGATCTCTGAGGCTTCAAGTGA 2007
Db 5158 ATGAAGCTCAGAGAGCTCAGAGAGGCCCTTTGCTTGGATCTCTGAGGCTTCAAGTGA 5099
QY 2008 TGTGATGCTTGGAGCCAGCACAACTCAAGCAAAATGACCAAGCCATGATATCTGACAG 2067
Db 5098 TGTGATGCTTGGAGCCAGCACAACTCAAGCAAAATGACCAAGCCATGATATCTGACAG 5039
QY 2068 ATTAATTAATGTTTGACCACTATTATGACCGCTGGAAGCAAGCAAGCAACAATTGCTC 2127
Db 5038 ATTAATTAATGTTTGACCACTATTATGATGATGCTGGAAGCAAGCAACAATTGCTC 4979
QY 2128 AACGTCCTCTCTGCGTGAATATGCTGTAAGTGGCTGCTGAATGTTATGATACGGGA 2187
Db 4978 AATGTCCTCTCTGCGTGAATATGCTGTAAGTGGCTGCTGAATGTTATGATACGGGA 4919
QY 2188 CGAAGAGGAGATCCGTGCTCTGCTTTAAACTGGCATCATTTCCCTGTGTAAGCA 2247
Db 4918 CGAAGAGGAGATCCGTGCTCTGCTTTAAACTGGCATCATTTCCCTGTGTAAGCA 4859
QY 2248 CATTTGAAGACAAAGTACAGATACCTTTCAAGCAAGTGGCAAGTCAAGAGATTTGT 2307
Db 4858 CACTTGAAGACAAAGTACAGATACCTTTCAAGCAAGTGGCAAGTCAAGAGATTTGT 4799
QY 2308 GACAGCGCAGGCTGGGCTCTCTGATGATCTAATCCAAATTCAGAGAGTGGGT 2367
Db 4798 GACAGCGTAGGCTGGGCTCTCTGATGATCTAATCCAAATTCAGAGAGTGGGT 4739
QY 2368 GAAGTTCATCTTTGGGGCAGTAACTTGAAGC 2401
Db 4738 GAAGTTCATCTTTGGGGCAGTAACTTGAAGC 4705

RESULT 4
US-09-427-048A-10/c
Sequence 10, Application US/09427048A
Patent No. 6203975
GENERAL INFORMATION:
APPLICANT: Trustees of the University of Pennsylvania
Wilson, James M.
Fisher, Krishna J.
Chen, Shu-Jen
Weltzman, Matthew
TITLE OF INVENTION: Improved Adenovirus Virus and
Methods of Use Thereof

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, P O Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/427,048A
FILING DATE: 21-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/836,022
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNVFN.008PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-427-048A-10

Query Match 62.8%; Score 1506.8; DB 3; Length 19307;
Best Local Similarity 90.6%; Pred. No. 0;
Matches 1607; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

| | | | |
|----|------|---|------|
| QY | 628 | TGGGCAACATCTGTAGATGACAGAACCGCTGGGTTCTTTACAGACACTCATAGA | 687 |
| DB | 6478 | TGGGGAACATTCATTAAGAGTAGTAGAGCAAGAGGCTGCTTTGGAAGAACTCATAGA | 6419 |
| QY | 688 | TTACTGCAACAGTCCCCCTGGACCTGGAAAAAGTTTCTTGCTGCTGCTTACAGAACTGAA | 747 |
| DB | 6418 | TTACTGCAACAGTCCCTCTGGACCTGGAGAAAGTTTCTTGCTGCTGCTTACCGAAGCAGAA | 6359 |
| QY | 748 | ACACTGCAATGTCTCTACAGATGCTACCCGTAAGGAAAGGCTCCTAGAGACTCCAG | 807 |
| DB | 6358 | ACACTGCAATGTCTCTACAGAGCGCTTCCCGTAAGGAGAGCTCCTAGAGACTCCAGG | 6299 |
| QY | 808 | GGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAGGTGAATGAAGCTCACACA | 867 |
| DB | 6298 | GGAGTCAAGAGCTGATGAACAATGGCAAGATCTCCAGAGAAATTGAAGCTCACACA | 6239 |
| QY | 868 | GATGTTATCAACACCTGATGAAAAACGCAAAATCTGAGATCCCTGGAAGGTTCC | 927 |
| DB | 6238 | GATATCTATCACAATCTTGATGAAATGGCCAAAAATCTGAGATCCCTGGAAGGTTCC | 6179 |
| QY | 928 | GATGATGAGTCTCTGTACAAAGAGCTTTGGATACATGAATTCAGTGAAGTGAAGTTC | 987 |
| DB | 6178 | GATGAGCAACCCCTGTACAAAGAGCTTTGGATACATGAATTCAGTGAAGTGAAGTTC | 6119 |
| QY | 988 | CGAAGAAAGTCTCTCAACATTAAGTCCCATTTTGAAGCAGTCTGACCAAGTGAAGCGT | 1047 |
| DB | 6118 | CAGAAAGAGTCTCTCAACATTAAGTCCCATTTTGAAGCAGTCTGACCAAGTGAAGCGT | 6059 |
| QY | 1048 | CTGACCTTTCTCTGACGAACTTCTGCTGTGCTACAGCTGAAGATGATGAATTAAGC | 1107 |
| DB | 6058 | TTGCACTTTCTCTCTGACGAACTTCTGCTGTGCTACAGCTGAAGATGATGAATTAAGC | 5999 |

| | | | |
|----|------|--|------|
| QY | 1108 | CGGACAGCACTATTGGAGGCGACTTTCACAGCACTTCAGAAAGCAAGATGTACATAGG | 1167 |
| DB | 5998 | CGTACAGCAACCATCGGTGTGATTTCCAGCAGCTTCAGAAAGCAAGATGTATATAGG | 5939 |
| QY | 1168 | GCCTTCAAGAGGAATTGAAAACTAAAGAACTGTATATGATGACTCTTGAGACTGTA | 1227 |
| DB | 5938 | GCCTTCAAGAGGAATTGAAAACTAAAGAACTGTATATGATGACTCTTGAGACTGTA | 5879 |
| QY | 1228 | CGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAACTCTACAGAGCCGAGA | 1287 |
| DB | 5878 | AGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAACTCTACAGAGCCGAGA | 5819 |
| QY | 1288 | GAGTGCCTCTGAGAGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG | 1347 |
| DB | 5818 | GAACTGCCTCTGAGAGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAA | 5759 |
| QY | 1348 | GAGTCAATACTGAGTGGAAAAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA | 1407 |
| DB | 5758 | GAGTCAACGCTGATGGACCAATTTGAACCTGCGCTCAGCTGATGGCAGAGAAAAATA | 5699 |
| QY | 1408 | GATGAGACCTTTGAAAGACTCCAGAACTTCAAGAGGCGCAAGATGAGCTGAGCTCAAG | 1467 |
| DB | 5698 | GATGAGCTCTTGAAGACTCCAGAACTTCAAGAGAGCTGCGCATGAATGAGCTCAAG | 5639 |
| QY | 1468 | CTGCGCCAAAGCTGAGGTGATCAAGGAACTCTGCGCAGCCCGTGGCGCATCTCCTCATTTAC | 1527 |
| DB | 5638 | TTGCGCCAAAGCTGAGGTGATCAAGGAACTCTGCGCAGCCCGTGGCGCATCTCCTCATTTAC | 5579 |
| QY | 1528 | TCTCTCCAGATCACTCTGAGAAAGTCAAGGCACTTCAAGAGAAATTTGCCCTCTGAAA | 1587 |
| DB | 5578 | TCTCTGCAAGATCACTTGAAGAAAGTCAAGGCACTTCCGGGAGAAATTTGCACCTCTTAAA | 5519 |
| QY | 1588 | GAGAACGTGAGCCACGCTCAATGACCTTGCTGCGCAGCTTACCACTTTGGGCAATTCAGCTC | 1647 |
| DB | 5518 | GAGATGCAATCGTGTCAATGACCTTGCAATGACCTTGCAATGACCTTGCGCAATTCAGCTC | 5459 |
| QY | 1648 | TCACCGTATAACCTCAGCACTCTGGAAGACCTGGAACACAGATGGAAGCTTCTGCAAGTG | 1707 |
| DB | 5458 | TCACCTTATAACCTCAGCACTTGAAGATCTGAATACAGATGAGGCTTCTGCAAGTG | 5399 |
| QY | 1708 | GCCGTGAGGACCGAGTCAAGGAGCTGCATGAAGCCCAAGGAACTTTGCTCAGCACTCT | 1767 |
| DB | 5398 | GCTGTGAGGACCGGTGTCAAGAGCTGCATGAAGCCCAAGGAACTTTGCTCAGCACTCT | 5339 |
| QY | 1768 | CAGCACTTCTTTCCAGCTCTGTCAGAGGCTCCCTGGAGAGAGCCATCTGCCAAACAA | 1827 |
| DB | 5338 | CAGCACTCTCTTCCAGCTCTGTCAGAGGCTCCCTGGAGAGAGCCATCTGCCAAACAA | 5279 |
| QY | 1828 | GTGCCCTACTATATCAACCAAGAGACTCAAAACAACTTGCTGGAGACCATCCCAAAATGACA | 1887 |
| DB | 5278 | GTGCCCTACTATATCAACCAAGAGACCCAAACCACTTGCTGGAGACCATCCCAAAATGACA | 5219 |
| QY | 1888 | GAGCTTACCAAGTCTTTAGCTGACCTGGAATATGTCAAGTTCTCAGCTTATAGACTGCC | 1947 |
| DB | 5218 | GAGCTTACCAAGTCTTTAGCTGACCTGGAATATGTCAAGTTCTCAGCTTATAGACTGCC | 5159 |
| QY | 1948 | ATGAAGTCCGAAGACTGCAAGAGGCTTTGCTTGATCTCTTGAGCCTGTCAAGTCCA | 2007 |
| DB | 5158 | ATGAAGTCAAGAGGCTCCAGAAAGGCTTTGCTTGATCTCTTGAGCCTGTCAAGTCCA | 5099 |
| QY | 2008 | TGTGATGCTTGGACCAAGCAACCTCAAGCAAAATGACCAAGCCCATGATATCTTGCAAG | 2067 |
| DB | 5098 | TGTGATGCTTGGACCAAGCAACCTCAAGCAAAATGACCAAGCCCATGATATCTTGCAAG | 5039 |
| QY | 2068 | ATTATTAATGTTTGAACCACTATTATGACCGCTGGAAGCAAGACACAAATTTGCTC | 2127 |
| DB | 5038 | ATAATTAATGTTTGAACCACTATTATGATGCTGCTGGAAGCAAGACACAAATTTGCTC | 4979 |
| QY | 2128 | AACGTCCTCTCTGCGTGATATGTCTGAACTGCTGCTGAATGTTATGATACGGGA | 2187 |
| DB | 4978 | AATGTCCTCTCTGCTGATGATGTCTCAACTGCTTCTCAATGTTATGATACGGGA | 4919 |
| QY | 2188 | CGAAGAGGAGATCCGTCCTGCTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCA | 2247 |

Db 4918 CGAACAGGAGGATCCGTGTCTCTCTTTAAACTGGCATCTTCTCTGTGTAAGCA 4859
QY 2248 CATTGGAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGAGATTGT 2307
Db 4858 CACTTGAAGACAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACTGGCTTGT 4799
QY 2308 GACCAGCGCAGGCTGGGCTCTCTGATGATTTCTATCCAAATTCACAGACAGTTGGGT 2367
Db 4798 GACCAGCGTAGCTGGGTCTTCTCTGATGATTTCTATCCAAATTCACAGACAGTTGGGT 4739
QY 2368 GAAGTTGCATCTTTGGGGGCAATACATTGAGC 2401
Db 4738 GAAGTTGCTTCTTTGGGGGCAATACATTGAGC 4705

RESULT 5
US-09-091-501B-7
Sequence 7, Application US/09091501B
Patent No. 6518413
GENERAL INFORMATION:
APPLICANT: Tinsley, Jonathon M
APPLICANT: Davies, Kay E
TITLE OF INVENTION: Ultraphin gene expression
FILE REFERENCE: 620-42
CURRENT APPLICATION NUMBER: US/09/091,501B
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: PCT/GB96/03156
PRIOR FILING DATE: 1996-12-19
PRIOR APPLICATION NUMBER: GB 9525962.8
PRIOR FILING DATE: 1995-12-19
PRIOR APPLICATION NUMBER: GB 9615797.9
PRIOR FILING DATE: 1996-07-26
PRIOR APPLICATION NUMBER: GB 9622174.2
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 6045
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: CDS
LOCATION: (11)..(6037)
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Chimeric
FEATURE:
NAME/KEY: misc feature
LOCATION: (724)..(758)
OTHER INFORMATION: Precise residue is left open
US-09-091-501B-7

Query Match 28.8%; Score 690.6; DB 4; Length 6045;
Best Local Similarity 60.5%; Pred. No. 1.1e-202;
Matches 1159; Conservative 0; Mismatches 749; Indels 9; Gaps 1;

QY 494 TGCTTCAAGAGATCTAGAACAGAGAGTCAAGGTCAATTTCTCACTCACATGTTG 553
Db 2871 TACTGCTTCAAGAGATGGGTCTCGAGATGGTATCGTCATGGCGTTGATAAGTCTGC 2930
QY 554 TGGTAGTGAATCTAGTGAAGTCAAGCACTGCTGCTTTGGAAGAACAACTTAAG 613
Db 2931 AGAACTCCTGAGAGATATGGAGTGAACACAGGAATGTGAAGAAACACAGAGT 2990
QY 614 TATTGGAGATGATGGGCAACATCTGTAGATGACAGAGACCGCTGGTTCTTTAC 673
Db 2991 ACTTAAACATCATGATCATCTCAACAAAGTATTGTGACAGACAGAGCGCTTG 3050
QY 674 AAGACACTCATAGATTACTGCAACAGTTCCCGCTGACCTGGAAGATTCTTGCCTGGC 733
Db 3051 AGGCTGAGTGAAGAGCGGTGACGCGCTCTCGAGAGATCTGGAACCTTCTGAGTGA 3110
QY 734 TTACAGAAAGCTGAACAACCTGCCAATGTCTTACAGGATGCTACCCGTAAGAAAGGCTCC 793

Db 3111 TCAGAGAGCAGAGACACAGATGAATGTGCTTGTGATGCTCTCATCGGAGATGCTC 3170
QY 794 TAGAAGCTCCAAAGGAGTAAGAGCTGTATGAACAAATGGCAAGACCTCCAAAGTAAA 853
Db 3171 TTCAGATAGTATCTTGGCCAGGAACTCAACAGCAGATGACAGATCCAGCAGAAA 3230
QY 854 TTGAAGCTCACAGATGTTTACAACTGATGAACCAAGCAAGCAAGCAAGCAAGT 913
Db 3231 TTGATGCCCAATGACATATTTAAAGCATTGACGGAAACAGAGAGATGTAAG 3290
QY 914 CCCTGAAGGTTCCGATGATGACGTCTGTTACAAAGAGCTTTGATAACATGAATTCA 973
Db 3291 CTTGGGAATTTCTGAAGAGGCTATGCTTCAACATGATGATGATGAACAAA 3350
QY 974 AGTGAAGTGAATTCGGAAGAAAGTCTCTCAACATTAGTCCATTGGAAGCCAGTTCTG 1033
Db 3351 GATGAATGACTTAAAGCAAAATCTGCTAGCATCAGGCGCCATTGGAGGCCAGCGCTG 3410
QY 1034 ACCAGTGAAGCGCTGACACCTTCTCTGACAGGAACCTTGTGCTACAGCTGAAG 1093
Db 3411 AGAAGTGAACAGGTTGCTGATGCTTGAAGAACTGATCAATGCTGAATATGAAG 3470
QY 1094 ATGATGAATTAAGCCGCGCAGGCACTATTGAGAGCGACTTTCCAGAGTTCAAGACAGA 1153
Db 3471 ATGAAGAGCTTAAGAAACAATGCTTATTGAGAGAGATGTTCCAGCCTTACAGCTCCAGT 3530
QY 1154 ACAGATGATAGGCGCTTCAAGAGGGAATGAAACCTTAAGAACTGTAATCATAGTA 1213
Db 3531 ATGACCATTTGAAGCCCTGAGAGGGAAGTTAAAGAGAAAGATATTCTGTCTGTAATG 3590
QY 1214 CTCTTGAGACTGTACGAATATTCTGACAGAGCAGCTTT-----GGAAGACTAG 1264
Db 3591 CTGTGACGAGCGCCGAGTTTCTTGGCTGATCAGCCAAATTGAGGCCCTGAAGAGCCAA 3650
QY 1265 AGAACTTAACAGAGAGCCCAAGAGCTGCTCTGAGAGAGAGCCCAAGATGTCACTC 1324
Db 3651 GAAGAACTTAACATCAAAAAGATTAATTAATCTCTGAGAGAGAGCCCAAGATGTCCA 3710
QY 1325 GGCTTCTAGAAAGCAGGCTGAGAGGTCAATAGTGGGAAAAATTGAACCTGCACT 1384
Db 3711 AAGCCATGCGCAACAGCTTTCTGAAGTCAAGAAAAATGGAAAGTCTAATGCTGTA 3770
QY 1385 CCGTGAAGTGAAGAAAAATAGTAGACCTTGAAGAACTCCAGAACTTCAAGAG 1444
Db 3771 CTAGCAATGGCAAAAAGCAAGTGAACAGCATTTGAAGAACTCAGAGACCTGACGGAG 3830
QY 1445 CCAGGATGAGCTGAGCTCAAGCTGCGCCAGCTGAGTCAAGGATCCTGGCAGC 1504
Db 3831 CTATGATGACCTGAGCGCTGACATGAAGAGCAGAGTCCGTGCGGAATGCTGAAGC 3890
QY 1505 CCGTGGCGAATCTCTCATTTGACTCTCTCAAGATCACTGAGAAAGTCAAGGCACTTC 1564
Db 3891 CCGTGGAGACTTAATCTCATTTGACTCTGACAGATCAATGAAGAAATCATGCGCATTTA 3950
QY 1565 GAGAGAAATTGGCGCTCTGAAGAGAACTGAGCCAGCTCAATGACCTTGCTGCGCAGC 1624
Db 3951 GAGAGAAATTGACCAATCACTTTAAAGTTAAACGGTGAATGATTATCCAGTCAAGC 4010
QY 1625 TTACCACTTTGGGCAATTCAGCTCTCAACCGTATACTCAAGCACTCTGGAAGACCTGAACA 1684
Db 4011 TGCTCCACTTGAACCTGATCCCTCTTAAAGATGTCTGCGCAGCTAGATGACCTTAATA 4070
QY 1685 CCAGATGAAGCTTCTGCAAGTGGCGCTGAGAACCGAGTCAAGCAGCTGATGAAGCCC 1744
Db 4071 TGGATGAAGACTTTTACAGGTTTCTGTGATGATGCGCCTTAACAGCTTCAAGGAAGCCC 4130
QY 1745 ACAGGACTTTGTTCCAGCATCTCAGCACTTTCTTCCAGCTGTGTTCCAGGATCCCTGGG 1804
Db 4131 ACAGAGATTTTGAACCATCTCTCAGCACTTTCTCTTACGTAAGTCCAGCTGCGCTGGC 4190
QY 1805 AGAGAGCCATCTGCGCAACAAAGTGCCTACTATATCAACCAAGAGACTCAACCAACTT 1864

db 4191 AAAGATCCATTTCACATATAAAGTGCCTATTACATCAACCATCAAAACACAGACCACCT 4250
2y GCTGGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAAATGTCA 1924
db GTTGGACCATCTTAAATGACCGAACTCTTTCAATCCCTTGCTGACCTGAATAATGTAC 4310
2y GATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGG 1984
db GTTTTCTGCCTACCGTACAGCAATCAAAATCCGAAGACTACAAAAGCACTATGTTGG 4370
2y ATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGACCAAGCACAACTCAAGCAAAAATG 2044
db ATCTCTTAGAGTTGAGTACACAAAATGAAATTTTCAAAACAGCACAAAGTTGAACCAAAATG 4430
2y ACCAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCACTATTTATGACCGCCTGG 2104
db ACCAGCTCCTCAGTTTCCAGATGTCACTCAACTGTCTGACAACTTATGATGACTTG 4490
2y AGCAAGAGCACAAATTTTGTCAACGTCCCTCTCTGCGTGGATATGTGTCTGAACTGGC 2164
db AGCAATGCATAAGGACCTGTCAACGTTCACCTCTGCTGATATGTGTCTCAATTGGT 4550
2y TGTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCTCTGCTTTTAAACTG 2224
db TGCTCAATGCTATGACACGGGTGAACTGGAAAAATTAGAGTGCAGAGTCTGAAGATTG 4610
2y GCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAAGTACAGATACCTTTTCAAGCAAG 2284
db GATTAATGCTCTCTCCAAAGGTCTCTTTGGAAGAAAAATACAGATATCTCTTTAAGGAAG 4670
2y TGGCAAGTTCAACAGGATTTGTGACCGAGCGGAGGATCCGTGTCTCTGCTGATGATTCTA 2344
db TTGCGGGCCGACAGAAATGTGTGACCAAGGAGCTGGGCTGTTTACTTTCATGATGCCA 4730
2y TCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGCAGTAACATTGAGC 2401
db TCCAGATCCCCCGCAGCTAGGTGAAGTAGCAGCTTTTGGAGGCAGTAAATATTGAGC 4787

RESULT 6

JS-09-091-501B-9
Sequence 9, Application US/09091501B
Patent No. 6518413

GENERAL INFORMATION:
APPLICANT: Tinsley, Jonathon M
APPLICANT: Davies, Kay E
TITLE OF INVENTION: Utrophin gene expression
FILE REFERENCE: 620-42
CURRENT APPLICATION NUMBER: US/09/091,501B
PRIOR FILING DATE: 1998-06-18
CURRENT APPLICATION NUMBER: PCT/GB96/03156
PRIOR FILING DATE: 1996-12-19
PRIOR APPLICATION NUMBER: GB 9525962.8
PRIOR FILING DATE: 1995-12-19
PRIOR APPLICATION NUMBER: GB 9615797.9
PRIOR FILING DATE: 1996-07-26
PRIOR APPLICATION NUMBER: GB 9622174.2
PRIOR FILING DATE: 1996-10-24
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9

LENGTH: 10320
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: CDS
LOCATION: (11)..(10312)
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Full length
OTHER INFORMATION: utrophin construct
FEATURE:
NAME/KEY: misc feature
LOCATION: (724)..(758)

OTHER INFORMATION: Precise residue is left open
US-09-091-501B-9
Query Match 28.8%; Score 690.6; DB 4; Length 10320;
Best Local Similarity 60.5%; Pred. No. 1.6e-202;
Matches 1159; Conservative 0; Mismatches 749; Indels 9; Gaps 1;
QY 494 TGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGG 553
Db 7146 TACTGCTTCAAGAACTGGGTCTCTGGAGATGGTATCGTTCATGGCGTTGATAAACGCTCTGC 7205
QY 554 TGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAAACAACTTAAGG 613
Db 7206 AGAAACTCCTGGAGGAATATGGGAGTGATGACACAAAGGAATGTGAAGAAACCAACAGAGT 7265
QY 614 TATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTAC 673
Db 7266 ACTTAAAAACATCATGGATCAATCTCAAAACAAAGTATGCTGACAGACAGAACGCTTGG 7325
QY 674 AAGACACTCATAGATTACTGCAACAGTTCCCTGACCTGGAAAAGTTTCTTGCTTGGC 733
Db 7326 AGGCTGAGTGGAGGACGGTGCAGGCCTCTCGCAGAGATCTGGAAAACCTTCTCTGAAGTGA 7385
QY 734 TTACAGAAAGCTGAAACAACTGCCAATGTCTTACAGATGCTACCCGTAAAGGAAAGGCTCC 793
Db 7386 TCCAAGAAAGCAGAGACCAAGTGAATGTGCTTGTGATGCTCTCATCGGAGAAATGCTC 7445
QY 794 TAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTTCCAAAGGTGAAA 853
Db 7446 TTCAAGGATAGTATCTTTGGCCAGGAACTCAAAACAGCAGATGCAGGACATCCAGGCAGAAA 7505
QY 854 TTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAAACAGCCAAAATCCTGAGAT 913
Db 7506 TTGATGCCCAATGACATATTTAAAGCATTTGACGGAACAGGCAGAGATGGTAAAG 7565
QY 914 CCCTGGAAGGTTCCGATGATGACGCTCTGTTTACAAAGACGTTTGGATAACATGAACCTCA 973
Db 7566 CTTTGGGAAATTTCTGAAGAGGCTACTATGCTTCAACATCGACTGGATGATATGAACCAA 7625
QY 974 AGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTGGAAGCCAGTTCTG 1033
Db 7626 GATGGAATGACTTAAAGCAAAATCTGCTAGCATCAGGGCCCATTTGGAGCCAGCGCTG 7685
QY 1034 ACCAGTGAAGCGTCTGCACCTTTCTCTGCGAGAACTTCTGCTGGTGTGGTACAGCTGAAA 1093
Db 7686 AGAAGTGAACAGGTTGCTGATGCTCTTAGAAGAACTGATCAAAATGGCTGAATATGAAAG 7745
QY 1094 ATGATGAATTAAGCCGCGAGGCACCTATTGGAGCGGACTTTCCAGAGTTCAGAGCAGAGA 1153
Db 7746 ATGAAGAGCTTAAGAAACAAATGCCTATTGGAGGAGATGTTCCAGCCTTACAGCTCCAGT 7805
QY 1154 ACGATGTACATAGGCGCTTCAAGAGGGAATTAAGAACTAAAGAACTGTGAATCATGAGTA 1213
Db 7806 ATGACCATTTGTAAGGCCCTGAGACGGGAGTTAAAGGAGAAAGAAATATTCTGCTCTGAATG 7865
QY 1214 CTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTT-----GGAAGGACTAG 1264
Db 7866 CTGTGACCGAGGCCGAGTTTCTTGGCTGATCAGCCCAATTGAGGCCCTTGAAGAGCCAA 7925
QY 1265 AGAAACTCTACAGAGCCCGAGAGCTGCTCTGAGGAGAGAGCCGAGAGATGTCACTC 1324
Db 7926 GAAGAAACCTACATCAAAACAGAAATTAACCTCTGAGGAGAGAGCCCAAAAGATTGCCA 7985
QY 1325 GGCTTCTACGAAAGCAGGCTGAGGAGGTCAATCTAGTGGGAAAAATTAACCTGCACT 1384
Db 7986 AAGCCATGCGCAACAGTCTTCTGAAGTCAAGAAAAATGGGAAAGTCTAAATGCTGTAA 8045
QY 1385 CCGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAGACTCCAGGAACTTCAAGAGG 1444
Db 8046 CTAGCAATTGGCAAAAGCAAGTGGACAAGGATTTGAGAACTCAGAGACCTGCAGGGAG 8105
QY 1445 CCACGGATGAGCTGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGC 1504


```
; Patent No. 6544786
; GENERAL INFORMATION:
; APPLICANT: Xiao, Xiao
; APPLICANT: Liu, Paul
; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPliced PE
; FILE REFERENCE: 00792
; CURRENT APPLICATION NUMBER: US/09/687,875A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/158,868
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 238
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pXX-C2 5' junction
US-09-687-875A-13
```

```
Query Match 2.6%; Score 63.6; DB 4; Length 238;
Best Local Similarity 94.3%; Pred. No. 1.1e-09;
Matches 66; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 924 TTCGATGATGAGTCCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGA 983
Db 169 TTCGACGACGCGAGTACTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGA 228

QY 984 ACTTCGGAAA 993
Db 229 ACTTCGGAAA 238
```

```
RESULT 12
US-09-668-313A-10
; Sequence 10, Application US/09668313A
; Patent No. 6503756
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Susan M. Freier
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF SYNTAXIN 4 INTERACTING PROTEIN EXPRESSION
; FILE REFERENCE: RTS-0127
; CURRENT APPLICATION NUMBER: US/09/668,313A
; CURRENT FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 247
; SEQ ID NO 10
; LENGTH: 2574
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (218)...(1891)
US-09-668-313A-10
```

```
Query Match 1.8%; Score 44.2; DB 4; Length 2574;
Best Local Similarity 54.7%; Pred. No. 0.0059;
Matches 88; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 1801 TGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCACGAGACTCAAACA 1860
Db 1733 TGGGAGGAAGCTTACACAGAGATGGAATCAAGTACTTTCATCAACCACGTCGACAGACC 1792

QY 1861 ACTTGCTGGGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAAT 1920
Db 1793 ACGTCTGGATCCACCCCGTGATGAGCGCCCTGAACCTGTCTGTGACAGAGGAGTGAA 1852

QY 1921 GTCAGATTCTCAGCTTATAGGACTGCCATGAACCTCCGAAG 1961
Db 1853 GAGGACTGTCCAGAGAGCTAACAGACCCGAAAGCTGATG 1893
```

RESULT 13

```
US-09-198-452A-1/c
; Sequence 1, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragment thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1
; LENGTH: 1230025
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(15000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (15001)..(30000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (30001)..(45000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (45001)..(60000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (60001)..(75000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (75001)..(90000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (90001)..(105000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (105001)..(120000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (120001)..(135000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (135001)..(150000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (150001)..(165000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (165001)..(180000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (180001)..(195000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (195001)..(210000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (210001)..(225000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (225001)..(240000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (240001)..(255000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (255001)..(270000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (270001)..(285000)
```

```
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (285001)..(300000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (300001)..(315000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (315001)..(330000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (330001)..(345000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (345001)..(360000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (360001)..(375000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (375001)..(390000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (390001)..(405000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (405001)..(420000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (420001)..(435000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (435001)..(450000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (450001)..(465000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (465001)..(480000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (480001)..(495000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (495001)..(510000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (510001)..(525000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (525001)..(540000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (540001)..(555000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (555001)..(570000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (570001)..(585000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (585001)..(600000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (600001)..(615000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (615001)..(630000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (630001)..(645000)
;/ OTHER INFORMATION: n=a or c or g or t
```

```
;/ NAME/KEY: misc_feature
;/ LOCATION: (645001)..(660000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (660001)..(675000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (675001)..(690000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (690001)..(705000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (705001)..(720000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (720001)..(735000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (735001)..(750000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (750001)..(765000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (765001)..(780000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (780001)..(795000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (795001)..(810000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (810001)..(825000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (825001)..(840000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (840001)..(855000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (855001)..(870000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (870001)..(885000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (885001)..(900000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
;/ LOCATION: (900001)..(915000)
;/ OTHER INFORMATION: n=a or c or g or t
;/ NAME/KEY: misc_feature
```

Query Match 1.8%; Score 44; DB 4; Length 1230025;
Best Local Similarity 53.5%; Pred. No. 0.47;

Matches 92; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```
QY 242 AATTATCAGAGATGAAGAACTGAAGTACAAGACGAGATGAATCTCCTAAATTCAGAT 301
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 656477 AATTAGAAGAAGAGAGAAAGAGAGAAATTTGAGGATATCAAAGACTCAGATACAAAT 656418
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 302 GGAATGCCTCAGGGTAGCTAGCATGGAACAAAGCAATTTACATAGAGTTTATGG 361
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 656417 GGGTTTCGATCACTCAAGCTGCTAAATTACATAACGTCACCTAGCAAGCAATTTATGTGG 656358
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 362 ATCTCCAGATCAGAAACTGAAGAGCTTGAATGACTGGCTAACAAAAACAGA 413
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 656357 CAATTAAGCAGAAAAAACTAAAGCTTCTTAAAGAGACGCGCTGGGAAATAGA 656306
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


RESULT 14

US-09-107-532A-1186
; Sequence 1186, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 1186:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1179 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...1179
; SEQUENCE DESCRIPTION: SEQ ID NO: 1186:
US-09-107-532A-1186
Query Match 1.8%; Score 43.4; DB 4; Length 1179;
Best Local Similarity 50.7%; Pred. No. 0.0061;
Matches 104; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
QY 371 ATCAGAACTGAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAACAAAGGAAAA 430
Db 602 AGCAAAGATAAAAGATTGTGATCGATCAGACAAAGAAAATGGAGATACGATCGGAGGAA 661
QY 431 TGGAGGAAGAGCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAAACACATA 490
Db 662 TTGTAGAAGTGTGCTGTGGAGCGTTTCCAGCTGGATTAGGAAGTACGTACAAATGGGACA 721
QY 491 AGGTGCTTCAAGAAGATCTAGAACACAGAACAAAGTCAGGGTCAATTCTCTCACTCATG 550
Db 722 CGAAGCTAGATGCCAAAATCGCACAAAGCTGTGGTTAGTATCAATGCCTTTAAAGGCGTAG 781
QY 551 TGGTGGTAGTTGATGAATCTAGTGG 575
Db 782 AATTGGGGTCGGATTCACTTCTGG 806

RESULT 15
US-09-620-312D-69
; Sequence 69, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; POLYPEPTIDES
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 69
; LENGTH: 1690
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (128)..(1522)
US-09-620-312D-69

Query Match 1.8%; Score 42.8; DB 4; Length 1690;
Best Local Similarity 47.9%; Pred. No. 0.012;
Matches 156; Conservative 0; Mismatches 167; Indels 3; Gaps 1;
QY 1428 CCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCCCAAGCTGAGGTGAT 1487
Db 55 CAAGGAGTTGCACCAGGTGGCGCAGCACCTGGACGAGAGCTGGCATGGTTTCAGGAGCG 114
QY 1488 CAAGGGA---TCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCT 1544
Db 115 GCTGCCACTGGCCATGCAGACAGAGCGAGGCAACGGTTTCAGGGCGGTCCAGCAGCACAT 174
QY 1545 CGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGT 1604
Db 175 CAAAAGAACAGGGCCTGCGGCGGAGATCCAGGGCGCATGGGCCCGCCCTGGAGGAGGT 234
QY 1605 CAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAG 1664
Db 235 GCTGGAGCGCGCGCGCTGGCGTGGCTGCGTGGCAGCCCGAGGAGGAGGTGCGCGG 294
QY 1665 CACTCTGGAAGACCTGAACACCAAGATGGAAGTTCTGAGGTGGCCGTCGAGGACCGAGT 1724
Db 295 GGGCCTTGAGCAGCTGCAGAGCGCTGGGCCGACTGCGGAGGCTGCCGAGGCGGCA 354
QY 1725 CAGGCAGCTGCATGAAGCCACAGGG 1750
Db 355 GCAGGTGTGGACGCCGCTTCCAGG 380

Search completed: February 2, 2004, 06:15:54

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run.on: February 1, 2004, 10:13:25 ; Search time 570.303 Seconds
(without alignments)
11364.749 Million cell updates/sec

Title: US-09-845-416-9_COPY_1000_3400

Perfect score: 2401

Sequence: 1 ggcagttcattgatggagag.....tgggggcagtaacattgagc 2401

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

- 1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
- 25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|-------|-------|--------|-------------|--------------------|
| 1 | 2401 | 100.0 | 3858 | 24 AAD37237 | Human dystrophin m |
| 2 | 2401 | 100.0 | 4825 | 24 AAD37257 | Adeno-associated v |
| 3 | 2401 | 100.0 | 4848 | 24 AAD37263 | Adeno-associated v |
| 4 | 2401 | 100.0 | 5060 | 24 AAD37254 | Adeno-associated v |
| 5 | 2250 | 93.7 | 3999 | 24 AAD37234 | Human dystrophin m |
| 6 | 2250 | 93.7 | 4966 | 24 AAD37256 | Adeno-associated v |
| 7 | 2250 | 93.7 | 4990 | 24 AAD37262 | Adeno-associated v |
| 8 | 2067 | 86.1 | 4182 | 24 AAD37230 | Human dystrophin m |

| | | | | | |
|----|--------|------|-------|-------------|--------------------|
| 9 | 2067 | 86.1 | 5149 | 24 AAD37255 | Adeno-associated v |
| 10 | 1839.4 | 76.6 | 5462 | 24 ABK81999 | DNA encoding mini- |
| 11 | 1737 | 72.3 | 3531 | 24 AAD37238 | Human dystrophin m |
| 12 | 1737 | 72.3 | 4498 | 24 AAD37258 | Adeno-associated v |
| 13 | 1735.6 | 72.3 | 5952 | 22 AAD06794 | Human dystrophin g |
| 14 | 1735.6 | 72.3 | 8689 | 24 ABK82000 | DNA encoding mini- |
| 15 | 1735.6 | 72.3 | 11058 | 24 AAD37229 | Human dystrophin p |
| 16 | 1735.6 | 72.3 | 11241 | 24 ABK82005 | cDNA encoding huma |
| 17 | 1735.6 | 72.3 | 11443 | 24 ABK82002 | DNA encoding mini- |
| 18 | 1735.6 | 72.3 | 12923 | 10 AAN90338 | Sequence of human |
| 19 | 1735.6 | 72.3 | 13957 | 24 ABT10904 | Human breast cance |
| 20 | 1735.6 | 72.3 | 13957 | 24 ABS69900 | Human dystrophin g |
| 21 | 1735.6 | 72.3 | 13957 | 24 ABN95786 | Gene #2284 used to |
| 22 | 1735.6 | 72.3 | 13957 | 24 ABK81959 | cDNA encoding huma |
| 23 | 1725 | 71.8 | 5339 | 24 ABK81998 | DNA encoding mini- |
| 24 | 1724.6 | 71.8 | 13977 | 24 ABS70403 | Human bone remodel |
| 25 | 1723 | 71.8 | 2169 | 24 AAD37232 | Human dystrophin r |
| 26 | 1695 | 70.6 | 3510 | 24 AAD37240 | Human dystrophin m |
| 27 | 1695 | 70.6 | 4476 | 24 AAD37259 | Adeno-associated v |
| 28 | 1506.8 | 62.8 | 13815 | 24 ABK81960 | cDNA encoding mous |
| 29 | 1506.8 | 62.8 | 13815 | 24 ABI99799 | Mouse ischaemic co |
| 30 | 1506.8 | 62.8 | 19307 | 17 AAT27558 | Shuttle vector pad |
| 31 | 1505.2 | 62.7 | 13815 | 19 AAV18885 | Mus musculus dyst |
| 32 | 1419 | 59.1 | 5417 | 24 ABK81997 | DNA encoding mini- |
| 33 | 1376 | 57.3 | 1821 | 24 AAD37241 | Human dystrophin r |
| 34 | 1263 | 52.6 | 4414 | 24 AAD37260 | Adeno-associated v |
| 35 | 1251 | 52.1 | 3446 | 24 AAD37242 | Human dystrophin m |
| 36 | 1058.2 | 44.1 | 4402 | 21 AAZ48567 | A rod shortened dy |
| 37 | 989 | 41.2 | 1434 | 24 AAD37243 | Human dystrophin r |
| 38 | 859 | 35.8 | 4402 | 21 AAZ48568 | A rod shortened dy |
| 39 | 723.4 | 30.1 | 4075 | 21 AAZ48569 | A rod shortened dy |
| 40 | 719.2 | 30.0 | 3747 | 21 AAZ48566 | A rod shortened dy |
| 41 | 708.2 | 29.5 | 11096 | 24 ABK81962 | cDNA encoding mous |
| 42 | 692.2 | 28.8 | 1991 | 24 AAD37231 | Human dystrophin N |
| 43 | 690.6 | 28.8 | 6045 | 18 AAT74665 | Utrophin minigene. |
| 44 | 690.6 | 28.8 | 6059 | 22 AAF84673 | Nucleotide sequenc |
| 45 | 690.6 | 28.8 | 10302 | 24 ABS69901 | Human utrophin gen |

ALIGNMENTS

RESULT 1

AAD37237
ID AAD37237 standard; DNA; 3858 BP.

XX AAD37237;

AC AAD37237;

XX 21-AUG-2002 (first entry)

DT Human dystrophin minigene delta3849.

DE Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

KW Becker muscular dystrophy; ds.

XX Homo sapiens.

XX WO200183695-A2.

XX 08-NOV-2001.

XX 27-APR-2001; 2001WO-US13677.

XX 28-APR-2000; 2000US-200777P.

XX (XIAO/) XIAO X.

XX Xiao X;

XX WPI; 2002-049342/06.

XX New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 48-49; 7lpp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3849 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1, R2), 8059-10227 (rods R22, R23 and R24, hinge H4
CC and CR domain) and 11047-11058 (dystrophin last 3 amino acids).
XX
SQ Sequence 3858 BP; 1189 A; 866 C; 905 G; 898 T; 0 other;

Query Match 100.0%; Score 2401; DB 24; Length 3858;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCAGTTCAATGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAA 60
Db |||||
QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAAT 120
Db |||||
QY 1060 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAAT 1119
QY 121 GATGTGGAAGTGGTGAAGACCAAGTTTCATCTCATGAGGGGTACATGATGGATTTGACA 180
Db |||||
QY 1120 GATGTGGAAGTGGTGAAGACCAAGTTTCATCTCATGAGGGGTACATGATGGATTTGACA 1179
QY 181 GCCCATCAGGCGCGGTTGGTAATATTCTACAATTTGGGAAGTAAAGCTGATTGGACAGGA 240
Db |||||
QY 1180 GCCCATCAGGCGCGGTTGGTAATATTCTACAATTTGGGAAGTAAAGCTGATTGGACAGGA 1239
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGA 300
Db |||||
QY 1240 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGA 1299
QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 360
Db |||||
QY 1300 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 1359
QY 361 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGA 420
Db |||||
QY 1360 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGA 1419
QY 421 ACAAGGAAATGGAGGAAGAGCCTCTTGGACCTGATCTTTGAAGACCTAAACGCCAAGTA 480
Db |||||
QY 1420 ACAAGGAAATGGAGGAAGAGCCTCTTGGACCTGATCTTTGAAGACCTAAACGCCAAGTA 1479
QY 481 CAACAACATAGGTGCTTCAAGAAGATCTAGAACAAAGACAGTCAAGGTCAATTCTCTC 540
Db |||||
QY 1480 CAACAACATAGGTGCTTCAAGAAGATCTAGAACAAAGACAGTCAAGGTCAATTCTCTC 1539
QY 541 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGAGATCACGCAACTGCTGCTTTGGAA 600
Db |||||
QY 1540 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGAGATCACGCAACTGCTGCTTTGGAA 1599
QY 601 GAACAACTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCGC 660
Db |||||
QY 1600 GAACAACTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCGC 1659
QY 661 TGGGTTCTTTTACAGACACTCATAGATTACTGCAACAGTTCCCTCGACCTGGAAAG 720
Db |||||
QY 1660 TGGGTTCTTTTACAGACACTCATAGATTACTGCAACAGTTCCCTCGACCTGGAAAG 1719
Db |||||

QY 721 TTTCTTGCCTGGCTTACAGAAGTGAACAACAACTGCCAATGTCTCTACAGGATGCTACCCGT 780
Db |||||
QY 1720 TTTCTTGCCTGGCTTACAGAAGTGAACAACAACTGCCAATGTCTCTACAGGATGCTACCCGT 1779
QY 781 AAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAGAGGTGATGAAACAATGGCAAGAC 840
Db |||||
QY 1780 AAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAGAGGTGATGAAACAATGGCAAGAC 1839
QY 841 CTCCAAGGTGAATTTGAAGCTCACACAGATGTTTTATCAACAACCTGGATGAAAACAGCCAA 900
Db |||||
QY 1840 CTCCAAGGTGAATTTGAAGCTCACACAGATGTTTTATCAACAACCTGGATGAAAACAGCCAA 1899
QY 901 AAAATCCTGAGATCCCTGGAAAGTTCGGATGATGCAGTCCCTGTTTACAAAAGACGTTGGAT 960
Db |||||
QY 1900 AAAATCCTGAGATCCCTGGAAAGTTCGGATGATGCAGTCCCTGTTTACAAAAGACGTTGGAT 1959
QY 961 AACATGAATCTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCATTTG 1020
Db |||||
QY 1960 AACATGAATCTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCATTTG 2019
QY 1021 GAAGCCAGTTCTGACCAAGTGGAAAGCTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGG 1080
Db |||||
QY 2020 GAAGCCAGTTCTGACCAAGTGGAAAGCTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGG 2079
QY 1081 CTACAGCTGAAAAGATGATGAATTAAGCCGGCAGGACCTATTTGGAGGCGACCTTTCCAGCA 1140
Db |||||
QY 2080 CTACAGCTGAAAAGATGATGAATTAAGCCGGCAGGACCTATTTGGAGGCGACCTTTCCAGCA 2139
QY 1141 GTTCAGAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTTGAAAACCTAAAGAACCT 1200
Db |||||
QY 2140 GTTCAGAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTTGAAAACCTAAAGAACCT 2199
QY 1201 GTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGCACAGAGCAGCCTTTTGGAAAGGA 1260
Db |||||
QY 2200 GTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGCACAGAGCAGCCTTTTGGAAAGGA 2259
QY 1261 CTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAGAGCCAGAAATGTC 1320
Db |||||
QY 2260 CTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAGAGCCAGAAATGTC 2319
QY 1321 ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATATCTGAGTGGGAAAAATTTGAACCTG 1380
Db |||||
QY 2320 ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATATCTGAGTGGGAAAAATTTGAACCTG 2379
QY 1381 CACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTCAA 1440
Db |||||
QY 2380 CACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTCAA 2439
QY 1441 GAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGG 1500
Db |||||
QY 2440 GAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGG 2499
QY 1501 CAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCA 1560
Db |||||
QY 2500 CAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCA 2559
QY 1561 CTTTCGAGGAGAAAATTGCGCCTCTGAAAGAGAAACGTGAGGCCACGTCAATGACCTTGTCTGC 1620
Db |||||
QY 2560 CTTTCGAGGAGAAAATTGCGCCTCTGAAAGAGAAACGTGAGGCCACGTCAATGACCTTGTCTGC 2619
QY 1621 CAGCTTACCACCTTTGGGCAATTCAGCTCTCACCCTATACCTCAGCACCTCTGGAAGACCTG 1680
Db |||||
QY 2620 CAGCTTACCACCTTTGGGCAATTCAGCTCTCACCCTATACCTCAGCACCTCTGGAAGACCTG 2679
QY 1681 AACACCAGATGGAAGCTTCTGAGGTGGCCCGTTCGAGGACCGAGTCAGGCAGCTGCATGAA 1740
Db |||||
QY 2680 AACACCAGATGGAAGCTTCTGAGGTGGCCCGTTCGAGGACCGAGTCAGGCAGCTGCATGAA 2739
QY 1741 GCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCC 1800
Db |||||
QY 2740 GCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCC 2799
Db |||||

QY 661 TGGGTTCTTTTACAAGACACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAG 720
Db 2417 TGGGTCTTTTACAAGACACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAG 2476
QY 721 TTTCTTGCCCTGGCTTACAGAAGCTGAAACAACACTGCCAATGTCTCTACAGGATGCTACCCGT 780
Db 2477 TTTCTTGCCCTGGCTTACAGAAGCTGAAACAACACTGCCAATGTCTCTACAGGATGCTACCCGT 2536
QY 781 AAGGAAAGGCTCCTAGAAGACTCCAAGGAGTAAAGAGCTGATGAAACAATGGCAAGAC 840
Db 2537 AAGGAAAGGCTCCTAGAAGACTCCAAGGAGTAAAGAGCTGATGAAACAATGGCAAGAC 2596
QY 841 CTCCAAGGTGAAATTGAAGCTCACACAGATGTTTTATCACAACCTGGATGAAAACAGCCAA 900
Db 2597 CTCCAAGGTGAAATTGAAGCTCACACAGATGTTTTATCACAACCTGGATGAAAACAGCCAA 2656
QY 901 AAAATCCTGAGATCCCTGGAAAGGTTCCGATGATGCAAGTCTCTTACAAAAGACGTTTGGAT 960
Db 2657 AAAATCCTGAGATCCCTGGAAAGGTTCCGATGATGCAAGTCTCTTACAAAAGACGTTTGGAT 2716
QY 961 AACATGAACCTTCAAGTGGAGTGAACTTCGGA AAAAGTCTCTCAACATTAGGTCCCATTTG 1020
Db 2717 AACATGAACCTTCAAGTGGAGTGAACTTCGGA AAAAGTCTCTCAACATTAGGTCCCATTTG 2776
QY 1021 GAAGCCAGTTCTGACCAGTGGAGCGGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGG 1080
Db 2777 GAAGCCAGTTCTGACCAGTGGAGCGGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGG 2836
QY 1081 CTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACTATTGGAGGCGACTTTCAGCA 1140
Db 2837 CTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACTATTGGAGGCGACTTTCAGCA 2896
QY 1141 GTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACCTAAAGAACCT 1200
Db 2897 GTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACCTAAAGAACCT 2956
QY 1201 GTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGAAGGA 1260
Db 2957 GTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGAAGGA 3016
QY 1261 CTAGAGAAACTCTACCAAGGAGCCAGAGAGCTGCCTCTGAGGAGAGAGCCAGAAATGTC 1320
Db 3017 CTAGAGAAACTCTACCAAGGAGCCAGAGAGCTGCCTCTGAGGAGAGAGCCAGAAATGTC 3076
QY 1321 ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAATTGAACCTG 1380
Db 3077 ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAATTGAACCTG 3136
QY 1381 CACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAA 1440
Db 3137 CACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAA 3196
QY 1441 GAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGG 1500
Db 3197 GAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGG 3256
QY 1501 CAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCA 1560
Db 3257 CAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCA 3316
QY 1561 CTTGAGAGAGAAATTGGCCCTCTGAAAGAGAACGTCGAGCCACGTCATGACCTTGCTCGC 1620
Db 3317 CTTGAGAGAGAAATTGGCCCTCTGAAAGAGAACGTCGAGCCACGTCATGACCTTGCTCGC 3376
QY 1621 CAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATACCTCAGCACTCTGGAAGACCTG 1680
Db 3377 CAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATACCTCAGCACTCTGGAAGACCTG 3436
QY 1681 AACACCAGATGGAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAA 1740
Db 3437 AACACCAGATGGAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAA 3496
QY 1741 GCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCCTGCCAGGGTCCC 1800

Db 3497 GCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTCCAGCTGTGCCAGGTCCC 3556
QY 1801 TGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACA 1860
Db 3557 TGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACA 3616
QY 1861 ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAAT 1920
Db 3617 ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAAT 3676
QY 1921 GTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGC 1980
Db 3677 GTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGC 3736
QY 1981 TTGATCTCTTGAGCCTGTGAGCTGTCATGTGATGCTGCTTGGACCAGCAACCTCAAGCAA 2040
Db 3737 TTGATCTCTTGAGCCTGTGAGCTGTCATGTGATGCTTGGACCAGCAACCTCAAGCAA 3796
QY 2041 AATGACCAGCCCATGGATATCTCTGCAGATTATTAATTGTTTGACCACATATTATGACCCG 2100
Db 3797 AATGACCAGCCCATGGATATCTCTGCAGATTATTAATTGTTTGACCACATATTATGACCCG 3856
QY 2101 CTGGAGCAAGAGCACACAATTTGGTCAACGTCCCTCTCTGCGTGGATATGTGCTGAAC 2160
Db 3857 CTGGAGCAAGAGCACACAATTTGGTCAACGTCCCTCTCTGCGTGGATATGTGCTGAAC 3916
QY 2161 TGGCTGCTGAATGTTTTATGATACGGGACGAAACAGGAGGATCCGTGTCTCTTTTAAA 2220
Db 3917 TGGCTGCTGAATGTTTTATGATACGGGACGAAACAGGAGGATCCGTGTCTCTTTTAAA 3976
QY 2221 ACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAGATACCTTTTCAAG 2280
Db 3977 ACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAGATACCTTTTCAAG 4036
QY 2281 CAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGAGGCTGGCCCTCTCTGTCATGAT 2340
Db 4037 CAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGAGGCTGGCCCTCTCTGTCATGAT 4096
QY 2341 TCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGATCCTTTGGGGGCAAGAACATTGAG 2400
Db 4097 TCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGATCCTTTGGGGGCAAGAACATTGAG 4156
QY 2401 C 2401
Db 4157 C 4157

RESULT 3
AAD37263
ID AAD37263 standard; DNA; 4848 BP.
XX

AC AAD37263;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3849.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Cytomegalovirus.
OS Chimeric - Unidentified.

XX WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.

Db 3460 AACACAGATGGAAGCTTTCAGAGTGGCCGTCGAGACCGAGTCAGGAGCTGCATGAA 3519
Qy 1741 GCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTCCACGTCGTCCAGGGTCCC 1800
Db 3520 GCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTCCACGTCGTCCAGGGTCCC 3579
Qy 1801 TGGGAGAGGCCATCTGCGCAAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 1860
Db 3580 TGGGAGAGGCCATCTGCGCAAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 3639
Qy 1861 ACTTGCTGGGACCATCCCAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAAT 1920
Db 3640 ACTTGCTGGGACCATCCCAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAAT 3699
Qy 1921 GTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAGGCCCTTTGC 1980
Db 3700 GTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAGGCCCTTTGC 3759
Qy 1981 TTGGATCTCTTGAGCCTGTGACGCTGCATGTGATGCCCTTGACGACGACACACCTCAAGCAA 2040
Db 3760 TTGGATCTCTTGAGCCTGTGACGCTGCATGTGATGCCCTTGACGACGACACACCTCAAGCAA 3819
Qy 2041 AATGACCAGCCCATGGATATCCTGTCAGATTATTAATTGTTTGACCACATAATTATGACCGC 2100
Db 3820 AATGACCAGCCCATGGATATCCTGTCAGATTATTAATTGTTTGACCACATAATTATGACCGC 3879
Qy 2101 CTGGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTGCGTGGATATGTCTGAAC 2160
Db 3880 CTGGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTGCGTGGATATGTCTGAAC 3939
Qy 2161 TGGTGTCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGTCTCTGTTTAAA 2220
Db 3940 TGGTGTCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGTCTCTGTTTAAA 3999
Qy 2221 ACTGGCATCATTTCCCTGTGTAAGCACATTTTGGAGACAAAGTACAGATACCTTTTCAAG 2280
Db 4000 ACTGGCATCATTTCCCTGTGTAAGCACATTTTGGAGACAAAGTACAGATACCTTTTCAAG 4059
Qy 2281 CAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGGAGGCTGGCCCTCTCTGTCATGAT 2340
Db 4060 CAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGGAGGCTGGCCCTCTCTGTCATGAT 4119
Qy 2341 TCTATCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGCAGTAACATTGAG 2400
Db 4120 TCTATCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGCAGTAACATTGAG 4179
Qy 2401 C 2401
Db 4180 C 4180

RESULT 4
ID AAD37264
XX AAD37264 standard; DNA; 5060 BP.
AC AAD37264;
XX
XX 21-AUG-2002 (first entry)
DT Adeno-associated virus (AAV) vector plasmid, AAV-E-CMV-3849.
DE
XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX Chimeric - Homo sapiens.
OS Chimeric - Cytomegalovirus.
OS Chimeric - Unidentified.
XX
PN WO200183695-A2.
XX 08-NOV-2001.

XX 27-APR-2001; 2001WO-US13677.
PF
XX 28-APR-2000; 2000US-200777P.
XX
XX (XIAO/) XIAO X.
PA
XX
XX Xiao X;
PI
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
PS
XX Example 1; Page 70-71; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC muscle creatine kinase (MCK) enhancer, a cytomegalovirus (CMV) promoter
CC and a small polyA signal sequence.
XX
SQ Sequence 5060 BP; 1449 A; 1217 C; 1234 G; 1160 T; 0 other;

Query Match 100.0%; Score 2401; DB 24; Length 5060;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA 60
Db 1992 GGAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA 2051
Qy 61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGCACAGGAGAGATTTCTAAT 120
Db 2052 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGCACAGGAGAGATTTCTAAT 2111
Qy 121 GATGTGAAGTGGTGAAGACACAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACA 180
Db 2112 GATGTGAAGTGGTGAAGACACAGTTTTCATCTACTCATGAGGGGTACATGATGGATTTGACA 2171
Qy 181 GCCATCAGGGCCGGTGTGTAATATTCTACAAATGGGAAGTAAGCTGATTTGAACAGGA 240
Db 2172 GCCATCAGGGCCGGTGTGTAATATTCTACAAATGGGAAGTAAGCTGATTTGAACAGGA 2231
Qy 241 AAATTATCAGAAGATGAAGAAACTGAAGTACAAAGACAGATGAATCTCTTAAATTTCAAGA 300
Db 2232 AAATTATCAGAAGATGAAGAAACTGAAGTACAAAGACAGATGAATCTCTTAAATTTCAAGA 2291
Qy 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 360
Db 2292 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 2351
Qy 361 GATCTCCAGAATCAGAAACTGAAAGAGTTTGAATGACTGGCTTAACAAAAACAGAAAGAA 420
Db 2352 GATCTCCAGAATCAGAAACTGAAAGAGTTTGAATGACTGGCTTAACAAAAACAGAAAGAA 2411
Qy 421 ACAAGGAAAAATGGAGGAAGAGCCTTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 480
Db 2412 ACAAGGAAAAATGGAGGAAGAGCCTTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 2471
Qy 481 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGTCAAGTCAAGGTCTCAATCTCTC 540
Db 2472 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGTCAAGGTCTCAATCTCTC 2531

QY 541 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAGCGAACTGCTGCTTTGGAA 600
DB 2532 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAGCGAACTGCTGCTTTGGAA 2591
QY 601 GAACAACCTTAAGGTATTGGGAGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCGC 660
DB 2592 GAACAACCTTAAGGTATTGGGAGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCGC 2651
QY 661 TGGGTTCTTTTACAAGACACCTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAG 720
DB 2652 TGGGTTCTTTTACAAGACACCTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAG 2711
QY 721 TTTCTTGCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGT 780
DB 2712 TTTCTTGCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGT 2771
QY 781 AAGGAAAGGCTCCTAGAAAGACTCCAAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGAC 840
DB 2772 AAGGAAAGGCTCCTAGAAAGACTCCAAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGAC 2831
QY 841 CTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGATAAACAGCCAA 900
DB 2832 CTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGATAAACAGCCAA 2891
QY 901 AAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCCTGTTACAAAGACGTTTGGAT 960
DB 2892 AAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCCTGTTACAAAGACGTTTGGAT 2951
QY 961 AACATGAATCAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAGGTCCCAATTG 1020
DB 2952 AACATGAATCAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAGGTCCCAATTG 3011
QY 1021 GAAGCCAGTCTTGACCAGTGGAAAGGCTCTGCACCTTTCTCTGCAGGAACCTTCTGCTGG 1080
DB 3012 GAAGCCAGTCTTGACCAGTGGAAAGGCTCTGCACCTTTCTCTGCAGGAACCTTCTGCTGG 3071
QY 1081 CTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGCA 1140
DB 3072 CTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGCA 3131
QY 1141 GTTCAGAAAGCAGAAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAACATAAGAACCT 1200
DB 3132 GTTCAGAAAGCAGAAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAACATAAGAACCT 3191
QY 1201 GTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGAAGGA 1260
DB 3192 GTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGAAGGA 3251
QY 1261 CTAGAGAAACTCTACCAGGAGCCACAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAAATGTC 1320
DB 3252 CTAGAGAAACTCTACCAGGAGCCACAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAAATGTC 3311
QY 1321 ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTG 1380
DB 3312 ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTG 3371
QY 1381 CACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAA 1440
DB 3372 CACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAA 3431
QY 1441 GAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGG 1500
DB 3432 GAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGG 3491
QY 1501 CAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCA 1560
DB 3492 CAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCA 3551
QY 1561 CTTTCGAGGAGAAAATTGCGCCTCTGAAAGAGAACGCTGAGCCACGTCATGACCTTGCTCGC 1620
DB 3552 CTTTCGAGGAGAAAATTGCGCCTCTGAAAGAGAACGCTGAGCCACGTCATGACCTTGCTCGC 3611

QY 1621 CAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAAACCTCAGCACTCTTGGAAAGACCTG 1680
DB 3612 CAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAAACCTCAGCACTCTTGGAAAGACCTG 3671
QY 1681 AACACCAGATGGAAGCTTCTGCGAGTGGCCGCTCGAGGACCGAGTCCAGGCAGCTGCATGAA 1740
DB 3672 AACACCAGATGGAAGCTTCTGCGAGTGGCCGCTCGAGGACCGAGTCCAGGCAGCTGCATGAA 3731
QY 1741 GCCCACAGGACTTTGGTCCAGCATCTCAGCACTTCTTCCACAGTCTGTCCAGGGTCCC 1800
DB 3732 GCCCACAGGACTTTGGTCCAGCATCTCAGCACTTCTTCCACAGTCTGTCCAGGGTCCC 3791
QY 1801 TGGGAGAGAGCCATCTCGCCAAAACAAAGTCCCCTACTATATCAACACGAGACTCAAAACA 1860
DB 3792 TGGGAGAGAGCCATCTCGCCAAAACAAAGTCCCCTACTATATCAACACGAGACTCAAAACA 3851
QY 1861 ACTTGCTGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAAT 1920
DB 3852 ACTTGCTGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAAT 3911
QY 1921 GTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGC 1980
DB 3912 GTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGC 3971
QY 1981 TTGGATCTCTTGAGCCTGTGAGCTGTCATGTGATGCCTTTGGACCAGCACAACTCAAGCAA 2040
DB 3972 TTGGATCTCTTGAGCCTGTGAGCTGTCATGTGATGCCTTTGGACCAGCACAACTCAAGCAA 4031
QY 2041 AATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCACCTATTTATGACCCG 2100
DB 4032 AATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCACCTATTTATGACCCG 4091
QY 2101 CTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGGATATGTGTCTGAAC 2160
DB 4092 CTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGGATATGTGTCTGAAC 4151
QY 2161 TGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGCTCTCTTTTAA 2220
DB 4152 TGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGCTCTCTTTTAA 4211
QY 2221 ACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAGATACCTTTTCAAG 2280
DB 4212 ACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAGATACCTTTTCAAG 4271
QY 2281 CAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGCGTGGGCCCTCCTTCTGCAATGAT 2340
DB 4272 CAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGCGTGGGCCCTCCTTCTGCAATGAT 4331
QY 2341 TCTATCCAAATTTCCAAGACAGTTGGGTGAAAGTTGATCCTTTGGGGCAGTAACATTTGAG 2400
DB 4332 TCTATCCAAATTTCCAAGACAGTTGGGTGAAAGTTGATCCTTTGGGGCAGTAACATTTGAG 4391
QY 2401 C 2401
DB 4392 C 4392
RESULT 5
AAD37234
ID AAD37234 standard; DNA; 3999 BP.
XX
AC AAD37234;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta3990.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX

PN WO200183695-A2.
XX 08-NOV-2001.
PD
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
XX (XIAO/) XIAO X.
PA
XX Xiao X;
PI
XX WPI; 2002-049342/06.
DR
XX
XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 46-47; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3990 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1 and R2), 8059-10227 (rods R22, R23 and R24, hinge
CC H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids).
XX
SQ Sequence 3999 BP; 1223 A; 907 C; 933 G; 936 T; 0 other;

Query Match 93.7%; Score 2250; DB 24; Length 3999;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 141; Gaps 1;

QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAA 60
DB 1000 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAA 1059

QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTAAT 120
DB 1060 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTAAT 1119

QY 121 GATGTGGAAGTGGTGAAAGACCAGTTTCATCTACTCATGAGGGGTACATGATGGATTGACA 180
DB 1120 GATGTGGAAGTGGTGAAAGACCAGTTTCATCTACTCATGAGGGGTACATGATGGATTGACA 1179

QY 181 GCCCATCAGGCGCGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGA 240
DB 1180 GCCCATCAGGCGCGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGA 1239

QY 241 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 300
DB 1240 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 1299

QY 301 TGGGAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 360
DB 1300 TGGGAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 1359

QY 361 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAGAAAGA 420
DB 1360 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAGAAAGA 1419

QY 421 ACAAGGAAAAATGAGGAAGAGCCCTCTTGGACCTGATCTTTGAAGACCTAAAAACGCCAAGTA 480
DB 1420 ACAAGGAAAAATGAGGAAGAGCCCTCTTGGACCTGATCTTTGAAGACCTAAAAACGCCAAGTA 1479

QY 481 CAACAAATAAGGTGCTTCAAGAAGATCTAGAAACAAGTCAAGGTCAATTCTCTC 540
DB 1480 CAACAAATAAGGTGCTTCAAGAAGATCTAGAAACAAGTCAAGGTCAATTCTCTC 1539

QY 541 ACTCATATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAGCGAACTGCTGCTTTGGAA 600
DB 1540 ACTCATATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAGCGAACTGCTGCTTTGGAA 1599

QY 601 GAACAACTTAAGGTATTGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCGC 660
DB 1600 GAACAACTTAAGGTATTGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCGC 1659

QY 661 TGGGTTCTTTTACAAGAC----- 678
DB 1660 TGGGTTCTTTTACAAGACCCAGCTGACCTAGCTCCTGGACTGACCACCTATTGGAGCCTCT 1719

QY 679 ----- 678
DB 1720 CCTACTCAGACTGTTACTCTGGTGACACAACTGTGGTTACTAAGGAAACTGCCATCTCC 1779

QY 679 -----ACTCATAGATTACTGCAACAG 699
DB 1780 AAAC TAGAAATGCCATCTTCCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 1839

QY 700 TTCCCCCTGGACCTGGAAGAGTTTCTTGCCCTGGCTTACAGAAGCTGAAACAACCTGCCAAT 759
DB 1840 TTCCCCCTGGACCTGGAAGAGTTTCTTGCCCTGGCTTACAGAAGCTGAAACAACCTGCCAAT 1899

QY 760 GTCTACAGGATGCTACCCGTAAGGAAAGGCTCTCTAGAAAGACTCCAAGGGAGTAAAGAG 819
DB 1900 GTCTACAGGATGCTACCCGTAAGGAAAGGCTCTCTAGAAAGACTCCAAGGGAGTAAAGAG 1959

QY 820 CTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAC 879
DB 1960 CTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAC 2019

QY 880 AACCTGATGAAACAGCCCAAAAATCCTGAGATCCTCGAAGGTTCCGATGATGCAGTC 939
DB 2020 AACCTGATGAAACAGCCCAAAAATCCTGAGATCCTCGAAGGTTCCGATGATGCAGTC 2079

QY 940 CTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACTTCGGAAGAGTCT 999
DB 2080 CTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACTTCGGAAGAGTCT 2139

QY 1000 CTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCT 1059
DB 2140 CTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCT 2199

QY 1060 CTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT 1119
DB 2200 CTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT 2259

QY 1120 ATTGGAGGCGACTTTCCAGCAAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGG 1179
DB 2260 ATTGGAGGCGACTTTCCAGCAAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGG 2319

QY 1180 GAATTGAAAACTAAAGAACCTTGAATCATGAGTACTCTTGAGACTGTACGAATATTCTG 1239
DB 2320 GAATTGAAAACTAAAGAACCTTGAATCATGAGTACTCTTGAGACTGTACGAATATTCTG 2379

QY 1240 ACAGAGCAGCCTTTTGGAAAGGACTAGAGAAAACCTCTACAGAGAGCCAGAGAGCTGCCTCT 1299
DB 2380 ACAGAGCAGCCTTTTGGAAAGGACTAGAGAAAACCTCTACAGAGAGCCAGAGAGCTGCCTCT 2439

QY 1300 GAGGAGAGAGCCAGAAATGTCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 1359
DB 2440 GAGGAGAGAGCCAGAAATGTCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 2499

QY 1360 GAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTT 1419
DB 2500 GAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTT 2559

QY 1420 GAAAGACTCCAGGAACTTCAAGAGGCGCCACGGATGAGCTGGACCTCAAGCTCGGCCAAGCT 1479

Db 2560 GAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGACCTCAAGCTGCGCCAAGCT 2619
2y 1480 GAGGTGATCAAGGGATCCTGGCAGCCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGAT 1539
2b 2620 GAGGTGATCAAGGGATCCTGGCAGCCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGAT 2679
2y 1540 CACCTCGAGAAAGTCAAGGCATCTCGAGGAGAAATTTGGCCTCTGAAAGAGAACGTGAGC 1599
2b 2680 CACCTCGAGAAAGTCAAGGCATCTCGAGGAGAAATTTGGCCTCTGAAAGAGAACGTGAGC 2739
2y 1600 CACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAAC 1659
Db 2740 CACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAAC 2799
2y 1660 CTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGCGAGGAC 1719
2b 2800 CTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGCGAGGAC 2859
2y 1720 CGAGTCAGGCAGCTGCATGAAGCCACAGGGACCTTTGGTCCAGCATCTCAGCACTTTCTT 1779
2b 2860 CGAGTCAGGCAGCTGCATGAAGCCACAGGGACCTTTGGTCCAGCATCTCAGCACTTTCTT 2919
2y 1780 TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT 1839
2b 2920 TCCAGCTGTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT 2979
2y 1840 ATCAACACAGAGACTCAAAACRACTTGTCTGGGACCATCCAAAATGACAGAGCTCTACCAG 1899
2b 2980 ATCAACACAGAGACTCAAAACRACTTGTCTGGGACCATCCAAAATGACAGAGCTCTACCAG 3039
2y 1900 TCTTTAGCTGACCTGAAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGA 1959
2b 3040 TCTTTAGCTGACCTGAAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGA 3099
2y 1960 AGACTGCAGAGAGGCCCTTTGCTGGATCTCTTGAGCCCTGTCAGCTGCATGATGACCTTG 2019
2b 3100 AGACTGCAGAGAGGCCCTTTGCTGGATCTCTTGAGCCCTGTCAGCTGCATGATGACCTTG 3159
2y 2020 GACCAGCAACCTCAAGCAAAATGACCCAGCCCATGGATATCCTGCAGATTATTAATTGT 2079
2b 3160 GACCAGCAACCTCAAGCAAAATGACCCAGCCCATGGATATCCTGCAGATTATTAATTGT 3219
2y 2080 TTGACCACTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTC 2139
2b 3220 TTGACCACTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTC 3279
2y 2140 TGGGTGGATATGTGTGTAAGTGGCTGCTGAATGTTTATGATACGGGACGACAGGGAGG 2199
Db 3280 TGGGTGGATATGTGTGTAAGTGGCTGCTGAATGTTTATGATACGGGACGACAGGGAGG 3339
2y 2200 ATCCGTGCTCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGAC 2259
2b 3340 ATCCGTGCTCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGAC 3399
2y 2260 AAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGCAGG 2319
2b 3400 AAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGCAGG 3459
2y 2320 CTGGGCTCCTTCTGTCATGATTCTATCCAAATTCOAAGACAGTTGGGTGAAGTTGCATCC 2379
2b 3460 CTGGGCTCCTTCTGTCATGATTCTATCCAAATTCOAAGACAGTTGGGTGAAGTTGCATCC 3519
2y 2380 TTTGGGGCAGTAACATTGAGC 2401
2b 3520 TTTGGGGCAGTAACATTGAGC 3541

RESULT 6
AAD37256
ID AAD37256 standard; DNA; 4966 BP.
XX
AC AAD37256;

XX 21-AUG-2002 (first entry)
DT Adeno-associated virus vector plasmid, AAV-MCK-delta3990.
XX
DE Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
KW
XX Chimeric - Homo sapiens.
OS Chimeric - Unidentified.
OS
XX WO200183695-A2.
FN 08-NOV-2001.
PD 27-APR-2001; 2001WO-US13677.
XX 28-APR-2000; 2000US-200777P.
PR (XIAO/) XIAO X.
XX Xiao X;
PA WPI; 2002-049342/06.
PI
XX
DR
XX
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 59-60; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence.
XX
SQ Sequence 4966 BP; 1403 A; 1216 C; 1232 G; 1115 T; 0 other;

Query Match 93.7%; Score 2250; DB 24; Length 4966;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 141; Gaps 1;

QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA 60
Db 1757 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA 1816
QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTCTTAAT 120
Db 1817 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTCTTAAT 1876
QY 121 GATGTGGAAGTGGTGAAGACCAGTTTCATCTACTCATGAGGGGTACATGATGGATTGACA 180
Db 1877 GATGTGGAAGTGGTGAAGACCAGTTTCATCTACTCATGAGGGGTACATGATGGATTGACA 1936
QY 181 GCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240
Db 1937 GCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 1996
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 300
Db 1997 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 2056
QY 301 TGGGAATGCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATG 360

Db 2057 TGGGAATGCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 2116
QY 361 GATCTCCAGAATCAGAAACTGAAGAGTTGAATGACTGGCTTAACAAAAACAGAGAAAGA 420
Db 2117 GATCTCCAGAATCAGAAACTGAAGAGTTGAATGACTGGCTTAACAAAAACAGAGAAAGA 2176
QY 421 ACAAGGAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTA 480
Db 2177 ACAAGGAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTA 2236
QY 481 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTC 540
Db 2237 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTC 2296
QY 541 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAGCGAACTGCTGCTTTGGAA 600
Db 2297 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAGCGAACTGCTGCTTTGGAA 2356
QY 601 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCGC 660
Db 2357 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCGC 2416
QY 661 TGGGTTCTTTTACAAGAC----- 678
Db 2417 TGGGTTCTTTTACAAGACCCAGCCTGACCTAGCTCCTGGACTGACCACCTATTGGAGCCTCT 2476
QY 679 ----- 678
Db 2477 CCTACTCAGACTGTTACTCTGTGTGACACAACCTGTGGTTACTAAGGAACTGCCATCTCC 2536
QY 679 -----ACTCATAGATTACTGCAACAG 699
Db 2537 AAAC TAGAATGCCATCTTCCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 2596
QY 700 TTCCCCCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACCTGCCAAT 759
Db 2597 TTCCCCCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACCTGCCAAT 2656
QY 760 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAAGCTCCAAGGGAGTAAAGAG 819
Db 2657 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAAGCTCCAAGGGAGTAAAGAG 2716
QY 820 CTGATGAACAATGGCAAGACCTCCAAGGTGAAATTGAAAGCTCACACAGATGTTTATCAC 879
Db 2717 CTGATGAACAATGGCAAGACCTCCAAGGTGAAATTGAAAGCTCACACAGATGTTTATCAC 2776
QY 880 AACCTGGATGAAACAGCCCAAAAATCCTGAGATCCCTGGAAAGTTCCGATGATGCAGTC 939
Db 2777 AACCTGGATGAAACAGCCCAAAAATCCTGAGATCCCTGGAAAGTTCCGATGATGCAGTC 2836
QY 940 CTGTTACAAGACGTTTTGGATAACATGAACCTTCAAGTGGAGTGAACTTCGGAAAAAGTCT 999
Db 2837 CTGTTACAAGACGTTTTGGATAACATGAACCTTCAAGTGGAGTGAACTTCCGAAAAAGTCT 2896
QY 1000 CTCAACATTAGGTCCCATTTTGGAGCCAGTTCTGACCAAGTGGAAAGCGTCTGCACTTTCT 1059
Db 2897 CTCAACATTAGGTCCCATTTTGGAGCCAGTTCTGACCAAGTGGAAAGCGTCTGCACTTTCT 2956
QY 1060 CTGCAGGAACCTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT 1119
Db 2957 CTGCAGGAACCTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT 3016
QY 1120 ATTGGAGGCGACTTTCCAGCAGTTTCAGAAGCAGAACCGATGATGATAGGGCCTTCAAGAGG 1179
Db 3017 ATTGGAGGCGACTTTCCAGCAGTTTCAGAAGCAGAACCGATGATGATAGGGCCTTCAAGAGG 3076
QY 1180 GAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 1239
Db 3077 GAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 3136
QY 1240 ACAGAGCGCCTTTTGAAGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCTCCT 1299

Db 3137 ACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCT 3196
QY 1300 GAGGAGAGAGCCCAGAAATGTCTACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 1359
Db 3197 GAGGAGAGAGCCCAGAAATGTCTACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 3256
QY 1360 GAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 1419
Db 3257 GAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 3316
QY 1420 GAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCT 1479
Db 3317 GAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCT 3376
QY 1480 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCTCATTTGACTCTCTCCAAGAT 1539
Db 3377 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCTCATTTGACTCTCTCCAAGAT 3436
QY 1540 CACCTCGAGAAAAAGTCAAGGCACCTCGAGGAGAAAAATTGGCCCTCTGAAAAGAGAACGTGAGC 1599
Db 3437 CACCTCGAGAAAAAGTCAAGGCACCTCGAGGAGAAAAATTGGCCCTCTGAAAAGAGAACGTGAGC 3496
QY 1600 CACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGSCATTCAGCTCTCACCGTATAAC 1659
Db 3497 CACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGSCATTCAGCTCTCACCGTATAAC 3556
QY 1660 CTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTTCAGGTTGCCGTCGAGGAC 1719
Db 3557 CTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTTCAGGTTGCCGTCGAGGAC 3616
QY 1720 CGAGTCAGGCAGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTT 1779
Db 3617 CGAGTCAGGCAGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTT 3676
QY 1780 TCCAGCTGTCTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTAT 1839
Db 3677 TCCAGCTGTCTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTAT 3736
QY 1840 ATCAACCAAGAGACTCAAAACAACCTTGTGGGACCATCCCAAAAATGACAGAGCTTACCAG 1899
Db 3737 ATCAACCAAGAGACTCAAAACAACCTTGTGGGACCATCCCAAAAATGACAGAGCTTACCAG 3796
QY 1900 TCTTTAGCTGACCTGAATAATGTGAGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGA 1959
Db 3797 TCTTTAGCTGACCTGAATAATGTGAGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGA 3856
QY 1960 AGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGATGATGCCTTG 2019
Db 3857 AGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGATGATGCCTTG 3916
QY 2020 GACCAAGCACAACCTCAAGCAAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGT 2079
Db 3917 GACCAAGCACAACCTCAAGCAAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGT 3976
QY 2080 TTGACCACTATTTATGACCCGCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTC 2139
Db 3977 TTGACCACTATTTATGACCCGCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTC 4036
QY 2140 TGCCTGGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGG 2199
Db 4037 TGCCTGGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGG 4096
QY 2200 ATCCGTCTCTGTCTTTTAAACTGGCATCATTTTCCCTGTGTAAAGCACATTTTGAAGAC 2259
Db 4097 ATCCGTCTCTGTCTTTTAAACTGGCATCATTTTCCCTGTGTAAAGCACATTTTGAAGAC 4156
QY 2260 AAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGGCAGG 2319
Db 4157 AAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGGCAGG 4216
QY 2320 CTGGGCTCCTTCTGCAATGATTTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGCATCC 2379
Db 4217 CTGGGCTCCTTCTGCAATGATTTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGCATCC 4276

QY 2380 TTTGGGGCAGTAACATTGAGC 2401
Db 4277 TTTGGGGCAGTAACATTGAGC 4298

RESULT 7

AAD37262
ID AAD37262 standard; DNA; 4990 BP.

XX AC AAD37262;

XX DT 21-AUG-2002 (first entry)

DE Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3990.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.

DS Chimeric - Homo sapiens.
DS Chimeric - Cytomegalovirus.
DS Chimeric - Unidentified.

XX WO200183695-A2.

XX 08-NOV-2001.

XX 27-APR-2001; 2001WO-US13677.

XX 28-APR-2000; 2000US-200777P.

XX (XIAO/) XIAO X.

XX Xiao X;

XX WPI; 2002-049342/06.

XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -

XX Example 1; Page 67-68; 71pp; English.

XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC cytomegalovirus (CMV) promoter and a small polyA signal sequence.

SQ Sequence 4990 BP; 1439 A; 1185 C; 1208 G; 1158 T; 0 other;

Query Match 93.7%; Score 2250; DB 24; Length 4990;

Best Local Similarity 94.5%; Pred. No. 0;

Matches 2401; Conservative 0; Mismatches 0; Indels 141; Gaps 1;

QY 1 GGCAGTTCAATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA 60

Db 1781 GGCAGTTCAATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA 1840

QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAAT 120

Db 1841 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAAT 1900

QY 121 GATGTGGAAGTGGTGAAGAACCCAGTTTCATCTACTGAGGGGTACATGATGGATTGACA 180

Db 1901 GATGTGGAAGTGGTGAAGAACCCAGTTTCATCTACTCATGAGGGGTACATGATGGATTGACA 1960
QY 181 GCCCATCAGGGCCGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGA 240
Db 1961 GCCCATCAGGGCCGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGA 2020
QY 241 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGCATGAATCTCTAAATTCAAGA 300
Db 2021 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGCATGAATCTCTAAATTCAAGA 2080
QY 301 TGGGAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATG 360
Db 2081 TGGGAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATG 2140
QY 361 GATCTCAGAATCAGAAACTGAAAGAGTTGAATGACTCGCTAACAAAAACAGAAAGAAGA 420
Db 2141 GATCTCAGAATCAGAAACTGAAAGAGTTGAATGACTCGCTAACAAAAACAGAAAGAAGA 2200
QY 421 ACAAGGAAAAATGGAGGAGAGCCTCTTGGACCTTGATCTTGAAGACCTAAAAACGCCAAGTA 480
Db 2201 ACAAGGAAAAATGGAGGAGAGCCTCTTGGACCTTGATCTTGAAGACCTAAAAACGCCAAGTA 2260
QY 481 CAACAACATAAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTCTCTC 540
Db 2261 CAACAACATAAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTCTCTC 2320
QY 541 ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 600
Db 2321 ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 2380
QY 601 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGC 660
Db 2381 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGC 2440
QY 661 TGGGTCTCTTTACAAGAC----- 678
Db 2441 TGGGTCTCTTTACAAGACCCAGCCTGACCTAGCTCCTCGACTGACCACCTATTGGAGCCTCT 2500
QY 679 ----- 678
Db 2501 CCTACTCAGACTGTTACTCTGGTGACACAACCTGTGGTTACTAAGGAAACTGCCATCTCC 2560
QY 679 -----ACTCATAGATTACTGCAACAG 699
Db 2561 AAACATAGAAATGCCATCTTCTCTGATGTTGGAGGTACTACTCATAGATTACTGCAACAG 2620
QY 700 TTCCCTCCTGGACCTGGAAAGTTTCTTGCTGCTTACAGAAAGCTGAAACAACCTGCCAAT 759
Db 2621 TTCCCTCCTGGACCTGGAAAGTTTCTTGCTGCTTACAGAAAGCTGAAACAACCTGCCAAT 2680
QY 760 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAGAG 819
Db 2681 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAGAG 2740
QY 820 CTGATGAACAACATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAC 879
Db 2741 CTGATGAACAACATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAC 2800
QY 880 AACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGAGTC 939
Db 2801 AACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGAGTC 2860
QY 940 CTGTTACAAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAANAAGTCT 999
Db 2861 CTGTTACAAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAANAAGTCT 2920
QY 1000 CTCACAATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAGCGTCTGCACCTTTCT 1059
Db 2921 CTCACAATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAGCGTCTGCACCTTTCT 2980
QY 1060 CTGCAGGAACCTTCTGGTGTGCTACAGCTGAAAGATGATGAATTAAAGCCGCAGGCACCT 1119

Db 2981 CTGAGGAACTTCTGGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT 3040

QY 1120 ATTGAGGCGACTTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGGGCCTTCAAGAGG 1179

Db 3041 ATTGAGGCGACTTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGGGCCTTCAAGAGG 3100

QY 1180 GAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTG 1239

Db 3101 GAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTG 3160

QY 1240 ACAGAGCAGCCTTTTGAAGGACTAGAGAAACTCTACAGAGAGCCAGAGAGCTGCCCTCT 1299

Db 3161 ACAGAGCAGCCTTTTGAAGGACTAGAGAAACTCTACAGAGAGCCAGAGAGCTGCCCTCT 3220

QY 1300 GAGGAGAGAGCCAGAAATGTCTACCTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACT 1359

Db 3221 GAGGAGAGAGCCAGAAATGTCTACCTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACT 3280

QY 1360 GAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCT 1419

Db 3281 GAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCT 3340

QY 1420 GAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCT 1479

Db 3341 GAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCT 3400

QY 1480 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGAT 1539

Db 3401 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGAT 3460

QY 1540 CACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGGCGCTCTGAAAGAGAACGTGAGC 1599

Db 3461 CACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGGCGCTCTGAAAGAGAACGTGAGC 3520

QY 1600 CACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAAC 1659

Db 3521 CACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAAC 3580

QY 1660 CTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTTGTCAGGTGGCCGTCGAGGAC 1719

Db 3581 CTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTTGTCAGGTGGCCGTCGAGGAC 3640

QY 1720 CGAGTCAGGAGCTGCATGAAGCCACACAGGGACTTTGGTCCAGCATCTCAGCACTTCTT 1779

Db 3641 CGAGTCAGGAGCTGCATGAAGCCACACAGGGACTTTGGTCCAGCATCTCAGCACTTCTT 3700

QY 1780 TCCACGTCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCAAAACAAAGTGCCCTACTAT 1839

Db 3701 TCCACGTCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCAAAACAAAGTGCCCTACTAT 3760

QY 1840 ATCAACCACGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACCAG 1899

Db 3761 ATCAACCACGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACCAG 3820

QY 1900 TCTTTAGCTGACCTGAATAATGTTCAGATTCTCAGCTTATAGACTGCCATGAAACTCCGA 1959

Db 3821 TCTTTAGCTGACCTGAATAATGTTCAGATTCTCAGCTTATAGACTGCCATGAAACTCCGA 3880

QY 1960 AGACTGCAGAGAGCCCTTTTGGCTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCTTG 2019

Db 3881 AGACTGCAGAGAGCCCTTTTGGCTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCTTG 3940

QY 2020 GACCAGCACACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTATTGT 2079

Db 3941 GACCAGCACACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTATTGT 4000

QY 2080 TTGACCACCTATTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTC 2139

Db 4001 TTGACCACCTATTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTC 4060

QY 2140 TCGGTGGATATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGACGAACAGGAGG 2199

Db 4061 TCGGTGGATATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGACGAACAGGAGG 4120

QY 2200 ATCCGTGTCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGAC 2259

Db 4121 ATCCGTGTCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGAC 4180

QY 2260 AAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGG 2319

Db 4181 AAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGG 4240

QY 2320 CTGGGCTCCTTCTGTCATCTTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCC 2379

Db 4241 CTGGGCTCCTTCTGTCATCTTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCC 4300

QY 2380 TTTGGGGCAGTAACATTGAGC 2401

Db 4301 TTTGGGGCAGTAACATTGAGC 4322

RESULT 8

AAD37230
ID AAD37230 standard; DNA; 4182 BP.

XX AAD37230;

AC AAD37230;

XX 21-AUG-2002 (first entry)

DT 21-AUG-2002 (first entry)

XX Human dystrophin minigene delta4173.

DE Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

KW Becker muscular dystrophy; ds.

OS Homo sapiens.

XX WO200183695-A2.

PN 08-NOV-2001.

XX 27-APR-2001; 2001WO-US13677.

PF 28-APR-2000; 2000US-200777P.

XX (XIAO/) XIAO X.

XX Xiao X;

XX WPI; 2002-049342/06.

DR New dystrophin minigene for treating Duchenne or Becker muscular

XX dystrophy comprises an N-terminal domain or modified N-terminal domain,

PT rod repeats, H1 and H4 domains and a cysteine rich domain of a

PT dystrophin gene -

XX Example 1; Page 43-44; 71pp; English.

PS The present invention relates to an isolated nucleotide sequence encoding

XX a dystrophin minigene. The minigene comprises N-terminal or modified

CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4

CC domains and cysteine-rich domains of dystrophin or utrophin genes. The

CC invention also relates to a recombinant adeno-associated virus (AAV)

CC comprising dystrophin minigene operably linked to an expression control

CC element. The dystrophin minigene in operable linkage with an expression

CC control element, in a recombinant adeno-associated virus or retrovirus is

CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular

CC dystrophy (BMD) in a mammalian subject. The present sequence is human

CC dystrophin minigene delta4173 containing nucleotides 1-1992 (N-terminus,

CC hinge H1 and rods R1, R2 and R3), 8059-10227 (rods R22, R23 and R24,

CC hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids).

XX Sequence 4182 BP; 1309 A; 927 C; 970 G; 976 T; 0 other;

SQ Query Match 86.1%; Score 2067; DB 24; Length 4182;

Best Local Similarity 88.1%; Pred. No. 0;

| | |
|---|--|
| Matches 2401; Conservative 0; Mismatches 0; Indels 324; Gaps 1; | |
| QY 1 | GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAA 60 |
| Db | |
| 1000 | GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAA 1059 |
| QY 61 | GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTTAAT 120 |
| Db | |
| 1060 | GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTTAAT 1119 |
| QY 121 | GATGTGGAAGTGGTGAAAGACCAGTTTCATCTACTCATGAGGGGTACATGATGGATTGACA 180 |
| Db | |
| 1120 | GATGTGGAAGTGGTGAAAGACCAGTTTCATCTACTCATGAGGGGTACATGATGGATTGACA 1179 |
| QY 181 | GCCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240 |
| Db | |
| 1180 | GCCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 1239 |
| QY 241 | AAATTATCAGAAGATGAAGAACTGAAGATCAAGAGCAGATGAATCTCCTAAATTCAAGA 300 |
| Db | |
| 1240 | AAATTATCAGAAGATGAAGAACTGAAGATCAAGAGCAGATGAATCTCCTAAATTCAAGA 1299 |
| QY 301 | TGGGAATGCCTCAGGGTAGCTAGCATGGAATAACAAAGCAATTTACATAGAGTTTAAATG 360 |
| Db | |
| 1300 | TGGGAATGCCTCAGGGTAGCTAGCATGGAATAACAAAGCAATTTACATAGAGTTTAAATG 1359 |
| QY 361 | GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAA 420 |
| Db | |
| 1360 | GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAA 1419 |
| QY 421 | ACAAGGAAAAATGGAGGAAGGCCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTA 480 |
| Db | |
| 1420 | ACAAGGAAAAATGGAGGAAGGCCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTA 1479 |
| QY 481 | CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTCTCTC 540 |
| Db | |
| 1480 | CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTCTCTC 1539 |
| QY 541 | ACTCAGATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 600 |
| Db | |
| 1540 | ACTCAGATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 1599 |
| QY 601 | GAACAACTTAAGGTATTGGGAGATCGATGGCAACAACATCTGTAGATGGACAGAAGACCGC 660 |
| Db | |
| 1600 | GAACAACTTAAGGTATTGGGAGATCGATGGCAACAACATCTGTAGATGGACAGAAGACCGC 1659 |
| QY 661 | TGGGTTCTTTTACAAGAC----- 678 |
| Db | |
| 1660 | TGGGTTCTTTTACAAGACATCCTTCTCAAAATGGCAACGTCTTACTGAAGAACAGTGCCTT 1719 |
| QY 679 | ----- 678 |
| Db | |
| 1720 | TTTATGTGATGGCTTTCAGAAAAAGAGATGCAGTGAACAAGATTTCACACAACCTGGCTTT 1779 |
| QY 679 | ----- 678 |
| Db | |
| 1780 | AAAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTA 1839 |
| QY 679 | ----- 678 |
| Db | |
| 1840 | GAAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACA 1899 |
| QY 679 | ----- 678 |
| Db | |
| 1900 | CTGAAGAAATAAGTCAGTGACCCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGT 1959 |
| QY 679 | ----- 678 |
| Db | |
| 1960 | TGGGATAAATTTAGTCCAAAACTTGAAAAAGAGTACAGCACAGACTCATAGATTACTGCAA 2019 |
| QY 697 | CAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCCTGGTTACAGAAGCTGAAACAACTGCC 756 |
| Db | |
| 2020 | CAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCCTGGTTACAGAAGCTGAAACAACTGCC 2079 |

| | |
|---------|---|
| QY 757 | AATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAGACTCCAAGGGAGTAAAA 816 |
| Db | |
| 2080 | AATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAGACTCCAAGGGAGTAAAA 2139 |
| QY 817 | GAGCTGATGAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTAT 876 |
| Db | |
| 2140 | GAGCTGATGAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTAT 2199 |
| QY 877 | CACAACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCA 936 |
| Db | |
| 2200 | CACAACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCA 2259 |
| QY 937 | GTCTGTGTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACTTCGGAAAAAG 996 |
| Db | |
| 2260 | GTCTGTGTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACTTCGGAAAAAG 2319 |
| QY 997 | TCTCTCAACATTAGGTCCCATTTGGAAAGCCAGTTCTGACAGTGGAAAGCGTCTGCACCTT 1056 |
| Db | |
| 2320 | TCTCTCAACATTAGGTCCCATTTGGAAAGCCAGTTCTGACAGTGGAAAGCGTCTGCACCTT 2379 |
| QY 1057 | TCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCA 1116 |
| Db | |
| 2380 | TCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCA 2439 |
| QY 1117 | CCTATTGGAGGCGACTTTCAGCAGATTCAAGAGCAGAACGATGTACATAGGGCCTTCAAG 1176 |
| Db | |
| 2440 | CCTATTGGAGGCGACTTTCAGCAGATTCAAGAGCAGAACGATGTACATAGGGCCTTCAAG 2499 |
| QY 1177 | AGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTT 1236 |
| Db | |
| 2500 | AGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTT 2559 |
| QY 1237 | CTGACAGAGCAGCCCTTTGGAAAGACTAGAGAACTCTACAGGAGCCCAGAGAGCTGCCT 1296 |
| Db | |
| 2560 | CTGACAGAGCAGCCCTTTGGAAAGACTAGAGAACTCTACAGGAGCCCAGAGAGCTGCCT 2619 |
| QY 1297 | CCTGAGGAGAGAGAGCCAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAAT 1356 |
| Db | |
| 2620 | CCTGAGGAGAGAGAGCCAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAAT 2679 |
| QY 1357 | ACTGAGTGGGAAAAATTGAACTTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGACC 1416 |
| Db | |
| 2680 | ACTGAGTGGGAAAAATTGAACTTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGACC 2739 |
| QY 1417 | CTTGAAAGACTCCAGGAACTTCAAGAGGCCACCGATGAGTGGACCTCAAGCTGCGCCAA 1476 |
| Db | |
| 2740 | CTTGAAAGACTCCAGGAACTTCAAGAGGCCACCGATGAGTGGACCTCAAGCTGCGCCAA 2799 |
| QY 1477 | GCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAA 1536 |
| Db | |
| 2800 | GCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAA 2859 |
| QY 1537 | GATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTG 1596 |
| Db | |
| 2860 | GATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTG 2919 |
| QY 1597 | AGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTAT 1656 |
| Db | |
| 2920 | AGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTAT 2979 |
| QY 1657 | AACCTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGCGCGTGAG 1716 |
| Db | |
| 2980 | AACCTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGCGCGTGAG 3039 |
| QY 1717 | GACCGAGTCAGGCAGCTGTCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGCACTTT 1776 |
| Db | |
| 3040 | GACCGAGTCAGGCAGCTGTCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGCACTTT 3099 |
| QY 1777 | CTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTAC 1836 |
| Db | |
| 3100 | CTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTAC 3159 |

QY 1837 TATATCAACACGAGACTCAAACTTGTCTGGGACATCCCAAATGACAGAGCTCTAC 1896
Db |||||||
3160 TATATCAACACGAGACTCAAACTTGTCTGGGACATCCCAAATGACAGAGCTCTAC 3219
QY 1897 CAGTCTTTAGCTGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAAACTC 1956
Db |||||||
3220 CAGTCTTTAGCTGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAAACTC 3279
QY 1957 CGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTTCAGCTGCATGTGATGCC 2016
Db |||||||
3280 CGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTTCAGCTGCATGTGATGCC 3339
QY 2017 TTGGACGAGCACAACTCAAGCAAAATGACCCCATGGATATCCTGCAGATTATTAAT 2076
Db |||||||
3340 TTGGACGAGCACAACTCAAGCAAAATGACCCCATGGATATCCTGCAGATTATTAAT 3399
QY 2077 TGTTTGACCACACTATTATGACCGCTGGAGCAAGAGCACAACTTTGGTCAACGTCCCT 2136
Db |||||||
3400 TGTTTGACCACACTATTATGACCGCTGGAGCAAGAGCACAACTTTGGTCAACGTCCCT 3459
QY 2137 CTCTGCGTGGATATGTCTGTAAGTGGCTGCTGAATGTTTATGATACGGGACGAACAGGG 2196
Db |||||||
3460 CTCTGCGTGGATATGTCTGTAAGTGGCTGCTGAATGTTTATGATACGGGACGAACAGGG 3519
QY 2197 AGGATCCGTGCTCCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAA 2256
Db |||||||
3520 AGGATCCGTGCTCCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAA 3579
QY 2257 GACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGC 2316
Db |||||||
3580 GACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGC 3639
QY 2317 AGGCTGGCCTCCTTCTGATGATTTCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCA 2376
Db |||||||
3640 AGGCTGGCCTCCTTCTGATGATTTCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCA 3699
QY 2377 TCCTTTGGGGGCAGTAACATTGAGC 2401
Db |||||||
3700 TCCTTTGGGGGCAGTAACATTGAGC 3724

RESULT 9
AAD37255
ID AAD37255 standard; DNA; 5149 BP.
XX
AC AAD37255;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-delta4173.
DE
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.

OS Chimeric - Homo sapiens.
OS Chimeric - Unidentified.
XX
FN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 57-59; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence.

XX
SQ Sequence 5149-BP; 1489 A; 1236 C; 1269 G; 1155 T; 0 other;

Query Match 86.1%; Score 2067; DB 24; Length 5149;
Best Local Similarity 88.1%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 324; Gaps 1;

QY 1 GGCAGTTCAATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA 60
Db |||||||
1757 GGCAGTTCAATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA 1816
QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTCGAAGCACAGGAGAGATTTCTAAT 120
Db |||||||
1817 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTCGAAGCACAGGAGAGATTTCTAAT 1876
QY 121 GATGTGGAAGTGGTGAAGACCAAGTTTTCATCTCATGAGGGGTACATGATGATTTGACA 180
Db |||||||
1877 GATGTGGAAGTGGTGAAGACCAAGTTTTCATCTCATGAGGGGTACATGATGATTTGACA 1936
QY 181 GCCCATCAGGGCCGGTGGTGAATAATTTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGA 240
Db |||||||
1937 GCCCATCAGGGCCGGTGGTGAATAATTTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGA 1996
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAGACAGATGAATCTCCTAAATTCAGA 300
Db |||||||
1997 AAATTATCAGAAGATGAAGAACTGAAGTACAAGACAGATGAATCTCCTAAATTCAGA 2056
QY 301 TGGGAATGCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATG 360
Db |||||||
2057 TGGGAATGCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATG 2116
QY 361 GATCTCCAGAATCAGAACTGAAGAGTTGAATGATGGCTTAACAAAAACAGAAAGAA 420
Db |||||||
2117 GATCTCCAGAATCAGAACTGAAGAGTTGAATGATGGCTTAACAAAAACAGAAAGAA 2176
QY 421 ACAAGGAAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTA 480
Db |||||||
2177 ACAAGGAAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTA 2236
QY 481 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAACAAGTCAGGGTCAATTCTCTC 540
Db |||||||
2237 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAACAAGTCAGGGTCAATTCTCTC 2296
QY 541 ACTCACATGGTGGTGTAGTTGATGAATCTAGTGAGATCAGCAACTGCTGCTTTGGAA 600
Db |||||||
2297 ACTCACATGGTGGTGTAGTTGATGAATCTAGTGAGATCAGCAACTGCTGCTTTGGAA 2356
QY 601 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACAATCTGTAGATGGACAGAACCCGC 660
Db |||||||
2357 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACAATCTGTAGATGGACAGAACCCGC 2416
QY 661 TGGGTTCTTTTACAAGAC----- 678
Db |||||||
2417 TGGGTTCTTTTACAAGACATCCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTT 2476

2Y 679 ----- 678
3b 2477 TTTAGTGCATGGCTTTTCAGAAAAAGAAAGATGCAGTGAACAAGATTACACAACTGGCTTT 2536
2Y 679 ----- 678
3b 2537 AAAGATCAAAAATGAAATGTTATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTA 2596
2Y 679 ----- 678
3b 2597 GAAAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACA 2656
2Y 679 ----- 678
Db 2657 CTGAAGAATAAGTCAGTGACCCAGAAGACGGAAGCATGGCTGGATAACTTTGGCCCGGTGT 2716
2Y 679 ----- ACTCATAGATTACTGCAA 696
3b 2717 TGGGATAATTAGTCCAAAAACTTGAAAAAGAGTACAGCACAGACTCATAGATTACTGCAA 2776
2Y 697 CAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCCTGCTTACAGAAGCTGAACAACAACTGCC 756
3b 2777 CAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCCTGGCTTACAGAAGCTGAACAACAACTGCC 2836
2Y 757 AATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAA 816
3b 2837 AATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAA 2896
2Y 817 GAGCTGATGAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTAT 876
3b 2897 GAGCTGATGAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTAT 2956
2Y 877 CACAACCTGGATGAAAAACAGCCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCA 936
3b 2957 CACAACCTGGATGAAAAACAGCCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCA 3016
2Y 937 GTCCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACTTCGGAAAAAG 996
3b 3017 GTCCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACTTCGGAAAAAG 3076
2Y 997 TCTCTCAACATTAGGTCCCATTTTGGAAAGCCAGTTCTGACAGTGGAAGCGTCTGCACCTT 1056
Db 3077 TCTCTCAACATTAGGTCCCATTTTGGAAAGCCAGTTCTGACAGTGGAAGCGTCTGCACCTT 3136
QY 1057 TCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCA 1116
Db 3137 TCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCA 3196
QY 1117 CCTATTGGAGCGGACTTTCCAGCAGTTTCCAGAGCAGAACGATGTACATAGGGCCTTCAAG 1176
Db 3197 CCTATTGGAGCGGACTTTCCAGCAGTTTCCAGAGCAGAACGATGTACATAGGGCCTTCAAG 3256
QY 1177 AGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAAATATT 1236
Db 3257 AGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAAATATT 3316
QY 1237 CTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACAGAGAGCCCAGAGAGCTGCCT 1296
Db 3317 CTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACAGAGAGCCCAGAGAGCTGCCT 3376
QY 1297 CCTGAGGAGAGAGCCCAAGATGTCTACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAAT 1356
Db 3377 CCTGAGGAGAGAGCCCAAGATGTCTACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAAT 3436
QY 1357 ACTGAGTGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACC 1416
Db 3437 ACTGAGTGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACC 3496
QY 1417 CTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCAA 1476
Db 3497 CTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCAA 3556
QY 1477 GCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAA 1536

Db 3557 GCTGAGGTGATCAAGGGATCCTTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAA 3616
QY 1537 GATCACTCGAGAAAGTCAAGSCACTTCGAGGAGAAATTCGCCCTCTGAAAGAGAACGTG 1596
Db 3617 GATCACTCGAGAAAGTCAAGSCACTTCGAGGAGAAATTCGCCCTCTGAAAGAGAACGTG 3676
QY 1597 AGCCACGTCAATGACCTTGTCCGCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTAT 1656
Db 3677 AGCCACGTCAATGACCTTGTCCGCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTAT 3736
QY 1657 AACCTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAGGTGGCCGTGAG 1716
Db 3737 AACCTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAGGTGGCCGTGAG 3796
QY 1717 GACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGACTTTGGTCCAGCATCTCAGCACTTT 1776
Db 3797 GACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGACTTTGGTCCAGCATCTCAGCACTTT 3856
QY 1777 CTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTAC 1836
Db 3857 CTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTAC 3916
QY 1837 TATATCAACCAAGAGACTCAAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTAC 1896
Db 3917 TATATCAACCAAGAGACTCAAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTAC 3976
QY 1897 CAGTCTTTAGCTGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTC 1956
Db 3977 CAGTCTTTAGCTGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTC 4036
QY 1957 CGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCTGTGAGCTGCATGTGATGCC 2016
Db 4037 CGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCTGTGAGCTGCATGTGATGCC 4096
QY 2017 TTGGACCAGCACAAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAAT 2076
Db 4097 TTGGACCAGCACAAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAAT 4156
QY 2077 TGTTTGACCACTATTTATGACCGCTGGAGCAAGAGACAAACAAATTTGGTCAACGTCCT 2136
Db 4157 TGTTTGACCACTATTTATGACCGCTGGAGCAAGAGACAAACAAATTTGGTCAACGTCCT 4216
QY 2137 CTCTGCGTGGATATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGG 2196
Db 4217 CTCTGCGTGGATATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGG 4276
QY 2197 AGGATCCGTGTCCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAA 2256
Db 4277 AGGATCCGTGTCCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAA 4336
QY 2257 GACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAGCGC 2316
Db 4337 GACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAGCGC 4396
QY 2317 AGGCTGGGCTCCTTCTGCAATGATTCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGCA 2376
Db 4397 AGGCTGGGCTCCTTCTGCAATGATTCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGCA 4456
QY 2377 TCCTTTGGGGCAGTAACATTGAGC 2401
Db 4457 TCCTTTGGGGCAGTAACATTGAGC 4481

RESULT 10

ABK81999
ID ABK81999 standard; DNA; 5462 BP.
XX
AC ABK81999;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding mini-dystrophin protein deltaR2-R21+H3.

XX Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.
KW

| | |
|----|---------------|
| OS | Homo sapiens. |
| OS | Synthetic. |

FN WO200229056-A2.

11-APR-2002.

AA
PF 04-OCT-2001; 2001WO-US31126.

PR 06-OCT-2000; 2000US-238848P.

PA (UNMI) UNIV MICHIGAN.

PI Chamberlain JS, Harper SQ;

DR WPI; 2002-435334/46.

A composition for preparing therapeutic drugs, has a mini-dystrophin peptide comprising a specific number of spectrin-like repeat domains, or a nucleic acid sequence encoding the mini-dystrophin peptide -

PS Disclosure; Fig 14; 145pp; English.

The invention describes a composition comprising a mini-dystrophin peptide comprising a spectrin-like repeat domain, where the domain comprises n spectrin-like repeats, and contains no more than n spectrin-like repeats, where n is an even number between 4-24, or a nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the polynucleotide encoding it is useful as a medicament, for preparing a drug for therapeutic application and in the preparation of a composition for treatment of muscle disease, e.g. Duchenne's muscular dystrophy (DMD). This sequence represents a mini-dystrophin sequence of the invention.

Sequence 5462 BP; 1668 A; 1225 C; 1212 G; 1357 T; 0 other;

```

Query Match          76.6%; Score 1839.4; DB 24; Length 5462;
Best Local Similarity 88.1%; Pred. No. 0;
Matches 2115; Conservative 0; Mismatches 76; Indels 210; Gaps 2;

```

Qy 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA 60

Db 1199 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA 1258

1199 GGCAGTTCAATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA 1258

QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTCTTAAT 120

Db 1259 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCCTAAT 1318

QY 121 GATGTGGAAGTGGTGAAGAGACCAGTTTCATACTCATGAGGGTACATGATCGATTGACA 180

D**b** 1319 GATGTGGAAGTGGTGAAAGACCAGTTTCATACTCATGAGGGGTACATGATGGATTGACA 1378

181 GCCCATCAGGGCCGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGA 240 QY

Db 1379 GCCCATCAGGGCCGGTTGGTAATATCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 1438

241 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGATCTCTAAATTC AAGA 300

D_b 1439 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 1498

301 TGGGAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTACATAGAGGTTTAAATG 360

db 1499 TGGGAA TGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATGCTCCTGGACTG 1558

361 GATCTCCAGAAATCAGAAAACCTGAAGAGATTGAATGACTGGCTAACAAAAACAGAGAAAGA 420

1558 ACA-----1562

421 ACAAGGAAAAATGGAGGAAGAGCCCTCTTTGGACCTTGATCTTTGAAGACCTAAACGCCAAGTA 480

| | | | |
|----|------|--|------|
| Db | 1563 | -----CTATTGGAGCCTCTCCTACTCAGACTGTTACTCTGGTG | 1600 |
| QY | 481 | CAACAACATAAGGTGCTTCAAGAAAGATCTAGAACAAGCAAGTCAGGGTCAATTCTCTC | 540 |
| Db | 1601 | ACACAACCTGTGGTTACTAAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTTCCCTTG | 1660 |
| QY | 541 | ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA | 600 |
| Db | 1661 | ATGTTGGAG----- | 1669 |
| QY | 601 | GAACAACCTTAAGGTATTGGGAGATCGATGGCAACAATCTGTAGATGGACAGAAAGACCGC | 660 |
| Db | 1670 | ----- | 1669 |
| QY | 661 | TGGGTTCTTTTACAAGACACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAG | 720 |
| Db | 1670 | -----CATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAG | 1708 |
| QY | 721 | TTTCTTGGCTGGCTTACAGAAGCTGAAAACAACTGCCAATGTCTCTACAGGATGCTACCCGT | 780 |
| Db | 1709 | TTTCTTGGCTGGCTTACAGAAGCTGAAAACAACTGCCAATGTCTCTACAGGATGCTACCCGT | 1768 |
| QY | 781 | AAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAAACAATGGCAAGAC | 840 |
| Db | 1769 | AAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAAACAATGGCAAGAC | 1828 |
| QY | 841 | CTCCAAGGTGAAAATTGAAGCTCACACAGATGTTTATCAACAACCTGGAATAAAACAGCCAA | 900 |
| Db | 1829 | CTCCAAGGTGAAAATTGAAGCTCACACAGATGTTTATCAACAACCTGGAATAAAACAGCCAA | 1888 |
| QY | 901 | AAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTGTACAAAGACGTTTGGAT | 960 |
| Db | 1889 | AAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTGTACAAAGACGTTTGGAT | 1948 |
| QY | 961 | AACATGAACCTCAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTG | 1020 |
| Db | 1949 | AACATGAACCTCAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTG | 2008 |
| QY | 1021 | GAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAACCTCTGGTGTGG | 1080 |
| Db | 2009 | GAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAACCTCTGGTGTGG | 2068 |
| QY | 1081 | CTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGCA | 1140 |
| Db | 2069 | CTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGCA | 2128 |
| QY | 1141 | GTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACATAAGAACCT | 1200 |
| Db | 2129 | GTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACATAAGAACCT | 2188 |
| QY | 1201 | GTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGA | 1260 |
| Db | 2189 | GTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGA | 2248 |
| QY | 1261 | CTAGAGAACTCTACAGGAGCCCAGAGAGCTGCCTCTCTGAGGAGAGAGCCCAGAAATGTC | 1320 |
| Db | 2249 | CTAGAGAACTCTACAGGAGCCCAGAGAGCTGCCTCTCTGAGGAGAGAGCCCAGAAATGTC | 2308 |
| QY | 1321 | ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTG | 1380 |
| Db | 2309 | ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTG | 2368 |
| QY | 1381 | CACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAAGACTCCAGGAACCTCAA | 1440 |
| Db | 2369 | CACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAAGACTCCAGGAACCTCAA | 2428 |
| QY | 1441 | GAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGG | 1500 |
| Db | 2429 | GAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGG | 2488 |
| QY | 1501 | CAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCA | 1560 |

2489 CAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCA 2548
1561 CTTTCGAGGAGAAATTGGCCCTCTGAAAAGAGAAAGTGGAGCCACGTCAATGACCTTGCTCGC 1620
2549 CTTTCGAGGAGAAATTGGCCCTCTGAAAAGAGAAAGTGGAGCCACGTCAATGACCTTGCTCGC 2608
1621 CAGCTTACCACCTTTTGGCAATTGAGCTCTCACCGTATTAACCTCAGCACTCTGGAAGACCTG 1680
2609 CAGCTTACCACCTTTTGGCAATTGAGCTCTCACCGTATTAACCTCAGCACTCTGGAAGACCTG 2668
1681 AACACCAGATGGAAGCTTCTGCAAGTGGCCGTGAGGACCGAGTCAGGAGCTGCATGAA 1740
2669 AACACCAGATGGAAGCTTCTGCAAGTGGCCGTGAGGACCGAGTCAGGAGCTGCATGAA 2728
1741 GCCCACAGGACCTTTGTGCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTTCCC 1800
2729 GCCCACAGGACCTTTGTGCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTTCCC 2788
1801 TGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCAGAGACTCAAACA 1860
2789 TGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCAGAGACTCAAACA 2848
1861 ACTTGCTGGGACCATCCCAAAATGACAGAGCTTACCAAGTCTTTTAGCTGACCTGAATAAT 1920
2849 ACTTGCTGGGACCATCCCAAAATGACAGAGCTTACCAAGTCTTTTAGCTGACCTGAATAAT 2908
1921 GTCAGATTCTCAGCTTATAGGACTGCCATGAACCTCCGAAGACTGCAGAAGGCCCTTTGC 1980
2909 GTCAGATTCTCAGCTTATAGGACTGCCATGAACCTCCGAAGACTGCAGAAGGCCCTTTGC 2968
1981 TTGGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCTTGGACCAAGCAACCTCAAGCAA 2040
2969 TTGGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCTTGGACCAAGCAACCTCAAGCAA 3028
2041 AATGACCAGCCCATGGATATCCTGCAGATTATTAATGTTTGGACCACTATTATGACCCG 2100
3029 AATGACCAGCCCATGGATATCCTGCAGATTATTAATGTTTGGACCACTATTATGACCCG 3088
2101 CTGGAGCAAGAGCACAAATTTGGTCAACGTCCCTCTCTGCGTGGATATGTCTGAAC 2160
3089 CTGGAGCAAGAGCACAAATTTGGTCAACGTCCCTCTCTGCGTGGATATGTCTGAAC 3148
2161 TGCGTGTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGCTCTTTTAAA 2220
3149 TGCGTGTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGCTCTTTTAAA 3208
2221 ACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAG 2280
3209 ACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAG 3268
2281 CAAGTGGCAAGTTCAACAGGATTTTGTGACCCAGCGCAGGCTGGGCCCTCTTCTGCATGAT 2340
3269 CAAGTGGCAAGTTCAACAGGATTTTGTGACCCAGCGCAGGCTGGGCCCTCTTCTGCATGAT 3328
2341 TCTATCCAAATTCOAAGACAGTTGGGTGAAGTTGCATCCTTTTGGGGGAGTAACATTGAG 2400
3329 TCTATCCAAATTCOAAGACAGTTGGGTGAAGTTGCATCCTTTTGGGGGAGTAACATTGAG 3388
2401 C 2401
3389 C 3389

RESULT 11
AAD37238
D AAD37238 standard; DNA; 3531 BP.

X AAD37238;
X
X
X 21-AUG-2002 (first entry)
X

Human dystrophin minigene delta3531.

KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
XX Becker muscular dystrophy; ds.
OS Homo sapiens.
XX WO200183695-A2.
PN 08-NOV-2001.
PD 27-APR-2001; 2001WO-US13677.
XX 28-APR-2000; 2000US-200777P.
PF (XIAO/) XIAO X.
PR XIAO X;
XX WPI; 2002-049342/06.
DR New dystrophin minigene for treating Duchenne or Becker muscular
XX dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX Example 1; Page 50-51; 71pp; English.
PS The present invention relates to an isolated nucleotide sequence encoding
XX a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3531 containing nucleotides 1-1341 (N-terminus,
CC hinge H1 and rod R1), 8059-10227 (rods R22, R23 and R24, hinge H4 and
CC CR domain) and 11047-11058 (dystrophin last 3 amino acids).
XX Sequence 3531 BP; 1071 A; 809 C; 824 G; 827 T; 0 other;
SQ Query Match 72.3%; Score 1737; DB 24; Length 3531;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 1765; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
Qy 576 AGATCAGCGAACTGCTGCTTTGGAGAAACAACCTTAAGGTATTGGGAGATCGATGGGCAA 635
Db 1248 AGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGGGAATG 1307
Qy 636 CATCTGTAGATGGACAGAAAGACCGCTGGGTCTTTTACAGACACTCATAGATTACTGCA 695
Db 1308 CCTCAGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAACTCATAGATTACTGCA 1367
Qy 696 ACAGTTCCCTCGACCTGGAAAAAGTTTCTTGCTGGCTTACAGAAAGCTGAAACAACTGC 755
Db 1368 ACAGTTCCCTCGACCTGGAAAAAGTTTCTTGCTGGCTTACAGAAAGCTGAAACAACTGC 1427
Qy 756 CAATGTCTACAGGATGCTACCCGTAAGGAAAGGCTCTTAGAAGACTCCAAGGGAGTAA 815
Db 1428 CAATGTCTACAGGATGCTACCCGTAAGGAAAGGCTCTTAGAAGACTCCAAGGGAGTAA 1487
Qy 816 AGAGCTGATGAACAAATGGCAAGACCTCCAAGGTGAAATGAAGCTCACACAGATGTTTA 875
Db 1488 AGAGCTGATGAACAAATGGCAAGACCTCCAAGGTGAAATGAAGCTCACACAGATGTTTA 1547
Qy 876 TCACAACTGGATGAAAAACAGCCAAAAATCCTCGAGATCCCTGGAAAGTTCCGATGATGC 935
Db 1548 TCACAACTGGATGAAAAACAGCCAAAAATCCTCGAGATCCCTGGAAAGTTCCGATGATGC 1607
Qy 936 AGTCTGTTACAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAA 995
XX

Db 1608 AGTCTGTTACAAAGACGTTTGGATAAATGAACTTCAAGTGGAGTGAACCTTCGGAAAAA 1667

QY 996 GTCTCTCAACATTAGGTCCCAATTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCT 1055

Db 1668 GTCTCTCAACATTAGGTCCCAATTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCT 1727

QY 1056 TTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGC 1115

Db 1728 TTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGC 1787

QY 1116 ACCTATTGGAGCGGACTTTCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAA 1175

Db 1788 ACCTATTGGAGCGGACTTTCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAA 1847

QY 1176 GAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATT 1235

Db 1848 GAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATT 1907

QY 1236 TCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCC 1295

Db 1908 TCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCC 1967

QY 1296 TCCTGAGGAGAGAGCCCCAGAAATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAA 1355

Db 1968 TCCTGAGGAGAGAGCCCCAGAAATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAA 2027

QY 1356 TACTGAGTGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGAC 1415

Db 2028 TACTGAGTGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGAC 2087

QY 1416 CCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGAACCTCAAGCTGCGCCA 1475

Db 2088 CCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGAACCTCAAGCTGCGCCA 2147

QY 1476 AGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCTCATTGACTCTCTCCA 1535

Db 2148 AGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCTCATTGACTCTCTCCA 2207

QY 1536 AGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAAATTGCGCTCTGAAAGAGAACGT 1595

Db 2208 AGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAAATTGCGCTCTGAAAGAGAACGT 2267

QY 1596 GAGCCACGTCAATGACCTTGCTCGCCAGTTACCACCTTTGGGCAATTCAGCTCTCACCGTA 1655

Db 2268 GAGCCACGTCAATGACCTTGCTCGCCAGTTACCACCTTTGGGCAATTCAGCTCTCACCGTA 2327

QY 1656 TAACCTCAGCACTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTGGCCGTGGA 1715

Db 2328 TAACCTCAGCACTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTGGCCGTGGA 2387

QY 1716 GGACCGAGTCAGGAGCTGCATGAAGCCACAGGGACTTTTGGTCCAGCATCTCAGCACTT 1775

Db 2388 GGACCGAGTCAGGAGCTGCATGAAGCCACAGGGACTTTTGGTCCAGCATCTCAGCACTT 2447

QY 1776 TCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCTTA 1835

Db 2448 TCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCTTA 2507

QY 1836 CTATATCAACCAGAGACTCAAAACAACCTTCTGTTGGACCATCCCAAAATGACAGAGCTCTA 1895

Db 2508 CTATATCAACCAGAGACTCAAAACAACCTTCTGTTGGACCATCCCAAAATGACAGAGCTCTA 2567

QY 1896 CCAGTCTTTAGTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACT 1955

Db 2568 CCAGTCTTTAGTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACT 2627

QY 1956 CCGAAGACTGCAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGC 2015

Db 2628 CCGAAGACTGCAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGC 2687

QY 2016 CTTGGACCAGCACACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAA 2075

Db 2688 CTTGGACCAGCACACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAA 2747

QY 2076 TTGTTTGACCACTATTTATGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCC 2135

Db 2748 TTGTTTGACCACTATTTATGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCC 2807

QY 2136 TCTCTGCGTGGATATGTGCTGAACTGGCTGCTGCTGAATGTTATGATACGGGACGACAGG 2195

Db 2808 TCTCTGCGTGGATATGTGCTGAACTGGCTGCTGCTGAATGTTATGATACGGGACGACAGG 2867

QY 2196 GAGGATCCGTGCTCTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTGGA 2255

Db 2868 GAGGATCCGTGCTCTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTGGA 2927

QY 2256 AGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAGCG 2315

Db 2928 AGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAGCG 2987

QY 2316 CAGGCTGGCCTCCTTCTGCAATGATTTCTATCCAAATTCGAAGACAGTTGGTGAAGTTGC 2375

Db 2988 CAGGCTGGCCTCCTTCTGCAATGATTTCTATCCAAATTCGAAGACAGTTGGTGAAGTTGC 3047

QY 2376 ATCCTTTGGGGCAGTAACATTGAGC 2401

Db 3048 ATCCTTTGGGGCAGTAACATTGAGC 3073

RESULT 12

AAD37258

ID AAD37258 standard; DNA; 4498 BP.

XX

AC AAD37258;

XX

DT 21-AUG-2002 (first entry)

XX

DE Adeno-associated virus vector plasmid, AAV-MCK-3531.

XX

KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin; adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD; Becker muscular dystrophy; ds.

KW

XX

OS Chimeric - Homo sapiens.

OS Chimeric - Unidentified.

XX

PN WO200183695-A2.

XX

PD 08-NOV-2001.

XX

PF 27-APR-2001; 2001WO-US13677.

XX

PR 28-APR-2000; 2000US-200777P.

XX

PA (XIAO/) XIAO X.

XX

PI Xiao X;

XX

DR WPI; 2002-049342/06.

XX

PT New dystrophin minigene for treating Duchenne or Becker muscular dystrophy comprises an N-terminal domain or modified N-terminal domain, rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene -

PS Example 1; Page 62-63; 71pp; English.

XX

CC The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular

QY 1708 GCCGTGAGGACCGAGTCAGGCAGCTGCGATGAAGCCACAGGGACTTTGGTCCAGCATCT 1767
Db |||||
QY 3982 GCCGTGAGGACCGAGTCAGGCAGCTGCGATGAAGCCACAGGGACTTTGGTCCAGCATCT 4041
Db |||||
QY 1768 CAGCACTTTCTTCCACGTCGTGCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 1827
Db |||||
QY 4042 CAGCACTTTCTTCCACGTCGTGCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 4101
Db |||||
QY 1828 GTGCCCTACTATATCAACACGAGACTCAACAACTTGCTGGGACCATCCCAAATGACA 1887
Db |||||
QY 4102 GTGCCCTACTATATCAACACGAGACTCAACAACTTGCTGGGACCATCCCAAATGACA 4161
Db |||||
QY 1888 GAGCTCTACCACTCTTTAGCTGACCTGAAATATGTCAGATTCTCAGCTTATAGGACTGCC 1947
Db |||||
QY 4162 GAGCTCTACCACTCTTTAGCTGACCTGAAATATGTCAGATTCTCAGCTTATAGGACTGCC 4221
Db |||||
QY 1948 ATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGCTGCA 2007
Db |||||
QY 4222 ATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGCTGCA 4281
Db |||||
QY 2008 TGTGATGCCCTTGACAGCAGCAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCGAG 2067
Db |||||
QY 4282 TGTGATGCCCTTGACAGCAGCAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCGAG 4341
Db |||||
QY 2068 ATTATTAATTGTTGACCACTATTTAAGACCGCCTGGAGCAAGAGCACAACAATTTGGTC 2127
Db |||||
QY 4342 ATTATTAATTGTTGACCACTATTTAAGACCGCCTGGAGCAAGAGCACAACAATTTGGTC 4401
Db |||||
QY 2128 AACGTCCCTCTCTGCGTGGATATGTGCTGAACTGGCTGCTGATGTTTATGATACGGGA 2187
Db |||||
QY 4402 AACGTCCCTCTCTGCGTGGATATGTGCTGAACTGGCTGCTGATGTTTATGATACGGGA 4461
Db |||||
QY 2188 CGAACAGGGAGGATCCGTGTCCTGCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCA 2247
Db |||||
QY 4462 CGAACAGGGAGGATCCGTGTCCTGCTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCA 4521
Db |||||
QY 2248 CATTTGGAAGACAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGT 2307
Db |||||
QY 4522 CATTTGGAAGACAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGT 4581
Db |||||
QY 2308 GACCAGCGCAGGCTGGGCCTCCTTCTGCAATGATTCTATCCAAATTTCAAGACAGTTGGST 2367
Db |||||
QY 4582 GACCAGCGCAGGCTGGGCCTCCTTCTGCAATGATTCTATCCAAATTTCAAGACAGTTGGST 4641
Db |||||

RESULT 14
ABK82000
ID ABK82000 standard; DNA; 8689 BP.

XX AC ABK82000;
XX XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding mini-dystrophin protein deltaH2-R19.
XX
KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200229056-A2.
XX
PD 11-APR-2002.
XX
XX 04-OCT-2001; 2001WO-US31126.
PF
XX 06-OCT-2000; 2000US-238848P.
PR

XX (UNMI) UNIV MICHIGAN.
PA Chamberlain JS, Harper SQ;
PI WPI; 2002-435334/46.
XX
DR A composition for preparing therapeutic drugs, has a mini-dystrophin
XX peptide comprising a specific number of spectrin-like repeat domains,
PT or a nucleic acid sequence encoding the mini-dystrophin peptide -
PT
XX Disclosure; Fig 15; 145pp; English.
XX
CC The invention describes a composition comprising a mini-dystrophin
CC peptide comprising a spectrin-like repeat domain, where the domain
CC comprises n spectrin-like repeats, and contains no more than n
CC spectrin-like repeats, where n is an even number between 4-24, or a
CC nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin
CC peptide or the polynucleotide encoding it is useful as a medicament,
CC for preparing a drug for therapeutic application and in the preparation
CC of a composition for treatment of muscle disease, e.g. Duchenne's
CC muscular dystrophy (DMD). This sequence represents a mini-dystrophin
CC sequence of the invention.
XX
SQ Sequence 8689 BP; 2721 A; 1804 C; 1861 G; 2303 T; 0 other;

Query Match 72.3%; Score 1735.6; DB 24; Length 8689;
Best Local Similarity 98.6%; Pred.No. 0;
Matches 1750; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 628 TGGGCAACATCTGTAGATGGACAGAGACCGCTGGTTCTTTTACAAGACACTCATAGA 687
Db |||||
QY 2948 TGGAGAGCATTATCAATAAGGGTGAGTGAGCGAGAGGCTGCTTTGGAAGAACTCATAGA 3007
Db |||||
QY 688 TTAGTGTCAACAGTTCCCTGGACCTGGAAAGATTTCTTGGCTTGGCTTACAGAAAGCTGAA 747
Db |||||
QY 3008 TTAGTGTCAACAGTTCCCTGGACCTGGAAAGATTTCTTGGCTTGGCTTACAGAAAGCTGAA 3067
Db |||||
QY 748 ACAACTGCCAATGTCTACAGGATGTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAG 807
Db |||||
QY 3068 ACAACTGCCAATGTCTACAGGATGTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAG 3127
Db |||||
QY 808 GGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAGTGAATTAAGCTCACACA 867
Db |||||
QY 3128 GGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAGTGAATTAAGCTCACACA 3187
Db |||||
QY 868 GATGTTTATCAACAACCTGGATGAACAAGCCCAAAAAATCCTGAGATCCCTGGAAAGTTCC 927
Db |||||
QY 3188 GATGTTTATCAACAACCTGGATGAACAAGCCCAAAAAATCCTGAGATCCCTGGAAAGTTCC 3247
Db |||||
QY 928 GATGATGCAGTCCTGTACAAAGACGTTTGGATGAACAATGAAGCTTCAAGTGGAGTGAACCTT 987
Db |||||
QY 3248 GATGATGCAGTCCTGTACAAAGACGTTTGGATGAACAATGAAGCTTCAAGTGGAGTGAACCTT 3307
Db |||||
QY 988 CGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCGT 1047
Db |||||
QY 3308 CGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCGT 3367
Db |||||
QY 1048 CTGCACCTTTCTCTCAGGAACCTTCTGGTGGCTACAGCTGAAAGATGATGAATTAAGC 1107
Db |||||
QY 3368 CTGCACCTTTCTCTCAGGAACCTTCTGGTGGCTACAGCTGAAAGATGATGAATTAAGC 3427
Db |||||
QY 1108 CGGAGGCACCTATTGGAGCGCATTTCAGCAGTTCAGAAAGCAGAACGATGTACATAGG 1167
Db |||||
QY 3428 CGGAGGCACCTATTGGAGCGCATTTCAGCAGTTCAGAAAGCAGAACGATGTACATAGG 3487
Db |||||
QY 1168 GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTA 1227
Db |||||
QY 3488 GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTA 3547
Db |||||
QY 1228 CGAATATTTCTGACAGAGCAGCCTTTTGGAAAGGACTAGAGAAACTTACCAGGAGCCCCAGA 1287
Db |||||
QY 3548 CGAATATTTCTGACAGAGCAGCCTTTTGGAAAGGACTAGAGAAACTTACCAGGAGCCCCAGA 3607
Db |||||

Db 8188 GGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAAGGTGAAATTGAAGTCAACA 8247
2Y 868 GATGTTTATCAAAACCTGGATGAAACAGCCAAAAATCCTGAGATCCCTGGAAGTTCC 927
Db 8248 GATGTTTATCAAAACCTGGATGAAACAGCCAAAAATCCTGAGATCCCTGGAAGTTCC 8307
2Y 928 GATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAAC TT 987
Db 8308 GATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAAC TT 8367
2Y 988 CGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCGT 1047
Db 8368 CGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCGT 8427
2Y 1048 CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 1107
Db 8428 CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 8487
2Y 1108 CGGCAGGCACCTATTGAGAGCGACTTCCAGCAGTTTCAAGAGCAGAAACGATGTACATAGG 1167
Db 8488 CGGCAGGCACCTATTGAGAGCGACTTCCAGCAGTTTCAAGAGCAGAAACGATGTACATAGG 8547
2Y 1168 GCCTTCAAGAGGGAATTGAAAACCTPAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTA 1227
Db 8548 GCCTTCAAGAGGGAATTGAAAACCTPAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTA 8607
2Y 1228 CGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAAACCTTACCCAGGAGCCCGAGA 1287
Db 8608 CGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAAACCTTACCCAGGAGCCCGAGA 8667
2Y 1288 GAGCTGCCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAAGCAGGCTGAG 1347
Db 8668 GAGCTGCCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAAGCAGGCTGAG 8727
2Y 1348 GAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACCTCGCTGACTGGCAGAGAAAAATA 1407
Db 8728 GAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACCTCGCTGACTGGCAGAGAAAAATA 8787
2Y 1408 GATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCGCACGGATGAGCTGGACCTCAAG 1467
Db 8788 GATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCGCACGGATGAGCTGGACCTCAAG 8847
2Y 1468 CTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCTCATTTGAC 1527
Db 8848 CTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCTCATTTGAC 8907
2Y 1528 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGGCGCTCTGAAA 1587
Db 8908 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGGCGCTCTGAAA 8967
2Y 1588 GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATCAGCTC 1647
Db 8968 GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATCAGCTC 9027
2Y 1648 TCACCGTATAACCTCAGCACCTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGAGGCTG 1707
Db 9028 TCACCGTATAACCTCAGCACCTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGAGGCTG 9087
2Y 1708 GCCGTCGAGGACCGAGTCAGGCAGCTGATGAAGCCACAGGGACTTTGGTCCAGCATCT 1767
Db 9088 GCCGTCGAGGACCGAGTCAGGCAGCTGATGAAGCCACAGGGACTTTGGTCCAGCATCT 9147
2Y 1768 CAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCCAACAAA 1827
Db 9148 CAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCCAACAAA 9207
2Y 1828 GTGCCCTACTATATCAACCAGGACTCAAAACACTTGTCTGGAGCCATCCCAAAATGACA 1887
Db 9208 GTGCCCTACTATATCAACCAGGACTCAAAACACTTGTCTGGAGCCATCCCAAAATGACA 9267
2Y 1888 GAGCTCTACCAAGTCTTTAGCTGACCTGAATAATGTCAAGATTCTCAGCTTATAGGACTGCC 1947

Db 9268 GAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAAGATTCTCAGCTTATAGGACTGCC 9327
QY 1948 ATGAAACTCCGAAGACTGCAAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAAGCTGCA 2007
Db 9328 ATGAAACTCCGAAGACTGCAAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAAGCTGCA 9387
QY 2008 TGTGATGCTCTTGACACGACCAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAG 2067
Db 9388 TGTGATGCTCTTGACACGACCAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAG 9447
QY 2068 ATTATTAATTGTTTGACCACTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTC 2127
Db 9448 ATTATTAATTGTTTGACCACTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTC 9507
QY 2128 AACGTCCCTCTCTGCGTGGATATGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGA 2187
Db 9508 AACGTCCCTCTCTGCGTGGATATGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGA 9567
QY 2188 CGAACAGGGAGGATCCGTGTCCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCA 2247
Db 9568 CGAACAGGGAGGATCCGTGTCCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCA 9627
QY 2248 CATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGT 2307
Db 9628 CATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGT 9687
QY 2308 GACCAGCGCAGGCTGGGCTCTCTTCTGCAATGATTTCTATCCAAATTTCCAAGACAGTTGGGT 2367
Db 9688 GACCAGCGCAGGCTGGGCTCTCTTCTGCAATGATTTCTATCCAAATTTCCAAGACAGTTGGGT 9747
QY 2368 GAAGTTGCATCCTTTGGGGGSCAGTAACATTGAGC 2401
Db 9748 GAAGTTGCATCCTTTGGGGGSCAGTAACATTGAGC 9781

Search completed: February 1, 2004, 12:04:30
Job time : 591.303 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 10:14:55 ; Search time 8553.13 Seconds
(without alignments)
11483.979 Million cell updates/sec

Title: US-09-845-416-9_COPY_1000_3400
Perfect score: 2401
Sequence: 1 ggcagttcattgatggagag.....tgggggcagtaacattgagc 2401

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB | ID | Description |
|------------|--------|---------------|--------|----|-----------|--------------------|
| 1 | 1839.4 | 76.6 | 5462 | 6 | AX538621 | AX538621 Sequence |
| 2 | 1735.6 | 72.3 | 5952 | 6 | AR304538 | AR304538 Sequence |
| 3 | 1735.6 | 72.3 | 5952 | 6 | AX114289 | AX114289 Sequence |
| 4 | 1735.6 | 72.3 | 8689 | 6 | AX538622 | AX538622 Sequence |
| 5 | 1735.6 | 72.3 | 11443 | 6 | AX538624 | AX538624 Sequence |
| 6 | 1735.6 | 72.3 | 12057 | 6 | AX538627 | AX538627 Sequence |
| 7 | 1735.6 | 72.3 | 12446 | 9 | HSDMDR | X14298 Human mRNA |
| 8 | 1735.6 | 72.3 | 13957 | 6 | AX409637 | AX409637 Sequence |
| 9 | 1735.6 | 72.3 | 13957 | 6 | AX538581 | AX538581 Sequence |
| 10 | 1735.6 | 72.3 | 13957 | 9 | HUMDYS | M18533 Homo sapien |
| 11 | 1725 | 71.8 | 5339 | 6 | AX538620 | AX538620 Sequence |
| 12 | 1724.6 | 71.8 | 13977 | 6 | AR220819 | AR220819 Sequence |
| 13 | 1540.4 | 64.2 | 13887 | 4 | AF070485 | AF070485 Canis fam |
| 14 | 1506.8 | 62.8 | 13815 | 6 | AX306153 | AX306153 Sequence |
| 15 | 1506.8 | 62.8 | 13815 | 6 | AX538582 | AX538582 Sequence |
| 16 | 1506.8 | 62.8 | 13815 | 10 | MUSDYSA | M68859 Mouse dystr |
| 17 | 1506.8 | 62.8 | 19307 | 6 | AR093392 | AR093392 Sequence |
| 18 | 1506.8 | 62.8 | 19307 | 6 | AR142592 | AR142592 Sequence |
| 19 | 1419 | 59.1 | 5417 | 6 | AX538619 | AX538619 Sequence |
| 20 | 1181.4 | 49.2 | 13575 | 5 | GGDYS | X13369 Chicken mrn |
| 21 | 1058.2 | 44.1 | 4402 | 6 | E30219 | E30219 Shortened d |
| 22 | 859 | 35.8 | 4402 | 6 | E30220 | E30220 Shortened d |
| 23 | 777.6 | 32.4 | 3521 | 5 | AF339031 | AF339031 Danio rer |
| 24 | 753 | 31.4 | 2654 | 5 | FSCDYSTRO | M37645 Torpedo cal |
| 25 | 723.4 | 30.1 | 4075 | 6 | E30221 | E30221 Shortened d |
| 26 | 719.2 | 30.0 | 3747 | 6 | E30218 | E30218 Shortened d |
| 27 | 711.2 | 29.6 | 10705 | 10 | RNAJ2967 | AJ002967 Rattus no |
| 28 | 708.2 | 29.5 | 11096 | 6 | AX538584 | AX538584 Sequence |
| 29 | 708.2 | 29.5 | 11096 | 10 | MMY12229 | Y12229 M.musculus |
| 30 | 690.6 | 28.8 | 6045 | 6 | A63605 | A63605 Sequence 7 |
| 31 | 690.6 | 28.8 | 6045 | 6 | AR281528 | AR281528 Sequence |
| 32 | 690.6 | 28.8 | 6059 | 6 | AX107972 | AX107972 Sequence |
| 33 | 690.6 | 28.8 | 10302 | 6 | AX538583 | AX538583 Sequence |
| 34 | 690.6 | 28.8 | 10302 | 9 | HSMUPS | X69086 H.sapiens m |
| 35 | 690.6 | 28.8 | 10320 | 6 | A63607 | A63607 Sequence 9 |
| 36 | 690.6 | 28.8 | 10320 | 6 | AR281529 | AR281529 Sequence |
| 37 | 671.6 | 28.0 | 3161 | 10 | MMGUTRPH | X83506 M.musculus |
| 38 | 665.6 | 27.7 | 3163 | 6 | E30223 | E30223 Shortened d |
| 39 | 576.6 | 24.0 | 3073 | 10 | AF195788 | AF195788 Rattus no |
| 40 | 576.6 | 24.0 | 3140 | 10 | AF195787 | AF195787 Rattus no |
| 41 | 576.6 | 24.0 | 3499 | 9 | HSU43519 | U43519 Human dystr |
| 42 | 576.6 | 24.0 | 5106 | 6 | AX552248 | AX552248 Sequence |
| 43 | 565.8 | 23.6 | 3275 | 10 | MUSDYS | M18025 Mouse dystr |
| 44 | 557 | 23.2 | 2110 | 9 | HUMDMXX | M92650 Human Duche |
| 45 | 557 | 23.2 | 4658 | 9 | BC028720 | BC028720 Homo sapi |

ALIGNMENTS

RESULT 1
AX538621
LOCUS AX538621 5462 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 41 from Patent WO0229056.
ACCESSION AX538621
VERSION AX538621.1 GI:25271168
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.

REFERENCE 1
AUTHORS Chamberlain,J.S. and Harper,S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 41 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

| | | | |
|----------------------------|-----------------------------|---|------------------------------------|
| FEATURES | | Location/Qualifiers | |
| source | 1.. 5462 | /organism="synthetic construct" | |
| | | /mol_type="genomic DNA" | |
| | | /db_xref="taxon:32630" | |
| | | /note="Synthetic" | |
| BASE COUNT | 1668 a 1225 c 1212 g 1357 t | | |
| ORIGIN | | | |
| Query Match | | 76.6%; | Score 1839.4; DB 6; Length 5462; |
| Best Local Similarity | | 88.1%; | Pred. No. 0; |
| Matches 2115; Conservative | | 0; | Mismatches 76; Indels 210; Gaps 2; |
| QY | 1 | GGCAGTTCATTGATGGAGAGTGAAGTAACCTGGACCGTTATCAACAGCTTTAGAA | 60 |
| Db | 1199 | GGCAGTTCATTGATGGAGAGTGAAGTAACCTGGACCGTTATCAACAGCTTTAGAA | 1258 |
| QY | 61 | GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAGGAGATTTCTAAT | 120 |
| Db | 1259 | GTATTATCGTGGCTTCTTCTGCTGAGACACATTGCAAGCACAGGAGATTTCTAAT | 1318 |
| QY | 121 | GATGTGGAAGTGGTGAAGACCAGTTTCATCTCATGAGGGTACATGATGGATTGACA | 180 |
| Db | 1319 | GATGTGGAAGTGGTGAAGACCAGTTTCATCTCATGAGGGTACATGATGGATTGACA | 1378 |
| QY | 181 | GCCCATCAGGCGCGGTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA | 240 |
| Db | 1379 | GCCCATCAGGCGCGGTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA | 1438 |
| QY | 241 | AAATTATCAGAAGATGAAGAAACTGAAGTACAAAGACAGATCAATCTCCTAAATTCAAGA | 300 |
| Db | 1439 | AAATTATCAGAAGATGAAGAAACTGAAGTACAAAGACAGATCAATCTCCTAAATTCAAGA | 1498 |
| QY | 301 | TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTACATAGAGTTTAAATG | 360 |
| Db | 1499 | TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTACATGCTCCTGGA | 1558 |
| QY | 361 | GATCTCCAGAACTCAGAAACTGAAAGAGTTGAATGACTGGCTAAACAAACAGAAAGAA | 420 |
| Db | 1559 | ACCA----- | 1562 |
| QY | 421 | ACAAGGAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTA | 480 |
| Db | 1563 | -----CTATTGGAGCCTCTCCTACTCAGACTGTTACTCTGGTG | 1600 |
| QY | 481 | CAACAACATAGGTGCTTCAAGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTCTCTC | 540 |
| Db | 1601 | ACACAACCTGTGTTACTAAGGAACTGCCATCTCCAAACTAGAAATGCCATCTTCTTG | 1660 |
| QY | 541 | ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA | 600 |
| Db | 1661 | ATGTTGGAG----- | 1669 |
| QY | 601 | GAACAACTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAAGCCGC | 660 |
| Db | 1670 | ----- | 1699 |
| QY | 661 | TGGGTTCTTTTACAAGACACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAG | 720 |
| Db | 1670 | -----CATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAG | 1708 |
| QY | 721 | TTTCTTGCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGT | 780 |
| Db | 1709 | TTTCTTGCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGT | 1768 |
| QY | 781 | AAGGAAAGGCTCCTPAGAAGACTCCAAGGAGTAAAGAGCTGATGAAACAAATGGCAAGAC | 840 |
| Db | 1769 | AAGGAAAGGCTCCTPAGAAGACTCCAAGGAGTAAAGAGCTGATGAAACAAATGGCAAGAC | 1828 |
| QY | 841 | CTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAACAGCCAA | 900 |
| Db | 1829 | CTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAACAGCCAA | 1888 |
| QY | 901 | AAAATCCTGAGATCCCTGGAAAGTTCGATGATGCAGTCCTGTTTACAAAAGACGTTTGGAT | 960 |
| Db | 1889 | AAAATCCTGAGATCCCTGGAAAGTTCGATGATGCAGTCCTGTTTACAAAAGACGTTTGGAT | 1948 |
| QY | 961 | AACATGAACCTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCATTTG | 1020 |
| Db | 1949 | AACATGAACCTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCATTTG | 2008 |
| QY | 1021 | GAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGG | 1080 |
| Db | 2009 | GAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGG | 2068 |
| QY | 1081 | CTACAGCTGAAAAGATGATGAATTAAGCCGGCAGGCACCTATTTGGAGGCGACTTTCAGCA | 1140 |
| Db | 2069 | CTACAGCTGAAAAGATGATGAATTAAGCCGGCAGGCACCTATTTGGAGGCGACTTTCAGCA | 2128 |
| QY | 1141 | GTTCAGAAACAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCT | 1200 |
| Db | 2129 | GTTCAGAAACAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCT | 2188 |
| QY | 1201 | GTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGCACAGACAGCCTTTTGGAAAGGA | 1260 |
| Db | 2189 | GTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGCACAGACAGCCTTTTGGAAAGGA | 2248 |
| QY | 1261 | CTAGAGAAACTCTACCAGGAGCCCGACAGAGCTGCCTCCTGAGGAGAGAGCCCGAATGTC | 1320 |
| Db | 2249 | CTAGAGAAACTCTACCAGGAGCCCGACAGAGCTGCCTCCTGAGGAGAGAGCCCGAATGTC | 2308 |
| QY | 1321 | ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGAACCTG | 1380 |
| Db | 2309 | ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGAACCTG | 2368 |
| QY | 1381 | CACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTCAA | 1440 |
| Db | 2369 | CACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTCAA | 2428 |
| QY | 1441 | GAGGCCACGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGG | 1500 |
| Db | 2429 | GAGGCCACGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGG | 2488 |
| QY | 1501 | CAGCCCGTGGCGGATCTCCTCATGTACTCTCTCCAAGATCACTCTCGAGAAAAGTCAAGCA | 1560 |
| Db | 2489 | CAGCCCGTGGCGGATCTCCTCATGTACTCTCTCCAAGATCACTCTCGAGAAAAGTCAAGCA | 2548 |
| QY | 1561 | CTTCGAGGAGAAAAATTGCGCCTCTGAAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGC | 1620 |
| Db | 2549 | CTTCGAGGAGAAAAATTGCGCCTCTGAAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGC | 2608 |
| QY | 1621 | CAGCTTACCACCTTTGGGCAATTGAGCTCTCACCGTATATAACCTCAGCACCTCTGGAAGACCTG | 1680 |
| Db | 2609 | CAGCTTACCACCTTTGGGCAATTGAGCTCTCACCGTATATAACCTCAGCACCTCTGGAAGACCTG | 2668 |
| QY | 1681 | AACACCAGATGGAAGCTTCTGCAAGTGGCCGTGAGGAGCCGAGTCAGGCAGCTGCATGAA | 1740 |
| Db | 2669 | AACACCAGATGGAAGCTTCTGCAAGTGGCCGTGAGGAGCCGAGTCAGGCAGCTGCATGAA | 2728 |
| QY | 1741 | GCCACACAGGACTTTTGGTCCAGCATCTCAGCACCTTTCTTCCACGTCTGTCCAGGTCCTCC | 1800 |
| Db | 2729 | GCCACACAGGACTTTTGGTCCAGCATCTCAGCACCTTTCTTCCACGTCTGTCCAGGTCCTCC | 2788 |
| QY | 1801 | TGGGAGAGAGCCCATCTCGCCCAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACA | 1860 |
| Db | 2789 | TGGGAGAGAGCCCATCTCGCCCAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACA | 2848 |
| QY | 1861 | ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAAT | 1920 |
| Db | 2849 | ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAAT | 2908 |
| QY | 1921 | GTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGC | 1980 |
| Db | 2909 | GTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGC | 2968 |
| QY | 1981 | TTGGATCTCTTGAGCCTGTCACTGCTGATGTGATGCCTTGGACCGACACAACTCAAGCAA | 2040 |

| | | | |
|----------------------------|------|--|------|
| Db | 2969 | TTGGATCTCTTGAGCCCTGTGAGCTGTCATGTGATGCCCTTGGACGAGCACAAACCTCAAGCAA | 3028 |
| Qy | 2041 | AATGACGAGCCCATGGATATCCTGCAGATTATTAATTTGTTGACCACTATTTATGACCGC | 2100 |
| Db | 3029 | AATGACGAGCCCATGGATATCCTGCAGATTATTAATTTGTTGACCACTATTTATGACCGC | 3088 |
| Qy | 2101 | CTGGAGCAAGAGACACAACAATTTGGTCAACGTCCTCTCTGCGTGGATATGTGCTGAAC | 2160 |
| Db | 3089 | CTGGAGCAAGAGACACAACAATTTGGTCAACGTCCTCTCTGCGTGGATATGTGCTGAAC | 3148 |
| Qy | 2161 | TGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCTCTGCTTTTAA | 2220 |
| Db | 3149 | TGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCTCTGCTTTTAA | 3208 |
| Qy | 2221 | ACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAAGTACAGATACCTTTTCAAG | 2280 |
| Db | 3209 | ACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAAGTACAGATACCTTTTCAAG | 3268 |
| 2y | 2281 | CAAGTGSAAAGTTCAACAGGATTTTGTGACCGAGCGCAGGCTGGCCCTCCTTCTGCAATGAT | 2340 |
| Db | 3269 | CAAGTGSAAAGTTCAACAGGATTTTGTGACCGAGCGCAGGCTGGCCCTCCTTCTGCAATGAT | 3328 |
| 2y | 2341 | TCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGCAGTAACATTGAG | 2400 |
| Db | 3329 | TCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGCAGTAACATTGAG | 3388 |
| 2y | 2401 | C 2401 | |
| Db | 3389 | C 3389 | |
| RESULT 2 | | | |
| AR304538 | | | |
| LOCUS | | | |
| DEFINITION | | | |
| ACCESSION | | | |
| VERSION | | | |
| KEYWORDS | | | |
| SOURCE | | | |
| ORGANISM | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| FEATURES | | | |
| source | | | |
| BASE COUNT | | | |
| ORIGIN | | | |
| Query Match | | | |
| Best Local Similarity | | | |
| Matches 1750; Conservative | | | |
| 2y | 628 | TGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACACTCATAGA | 687 |
| Db | 2902 | TGGAGAACCATTCATAAAGGGTGAAGTGAAGGCTGCTTTTGGAAAGAACTCATAGA | 2961 |
| 2y | 688 | TTACTGCAACAGTTTCCCTGGACCTGGAAAAAGTTTCTTGCCTGGCTTACAGAAAGCTGAA | 747 |
| Db | 2962 | TTACTGCAACAGTTTCCCTGGACCTGGAAAAAGTTTCTTGCCTGGCTTACAGAAAGCTGAA | 3021 |
| 2y | 748 | AACTGCAATGTCTCTACAGGATGCTACCCGTAAGGAAGGCTCCTAGAAGACTCCAAG | 807 |
| Db | 3022 | AACTGCAATGTCTCTACAGGATGCTACCCGTAAGGAAGGCTCCTAGAAGACTCCAAG | 3081 |
| 2y | 808 | GGAGTAAAGAGCTGTATGAAACAAATGGCAAGACCTCCAGGTGAAATTTGAAGCTCACACA | 867 |
| Db | 3082 | GGAGTAAAGAGCTGTATGAAACAAATGGCAAGACCTCCAGGTGAAATTTGAAGCTCACACA | 3141 |

| | | | |
|----|------|--|------|
| Qy | 868 | GATGTTTATCAAAACCTGGATGAAAAACAGCCAAAAATCCTGAGATCCCTGGAAGGTTCC | 927 |
| Db | 3142 | GATGTTTATCAAAACCTGGATGAAAAACAGCCAAAAATCCTGAGATCCCTGGAAGGTTCC | 3201 |
| Qy | 928 | GATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCT | 987 |
| Db | 3202 | GATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCT | 3261 |
| Qy | 988 | CGGAAAAAGTCTCTCAACATTAGTCCCATTTTGAAGCCAGTTTCTGACCAAGTGAAGCGT | 1047 |
| Db | 3262 | CGGAAAAAGTCTCTCAACATTAGTCCCATTTTGAAGCCAGTTTCTGACCAAGTGAAGCGT | 3321 |
| Qy | 1048 | CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC | 1107 |
| Db | 3322 | CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC | 3381 |
| Qy | 1108 | CGGCAGGACCTATTGGAGGCGACTTTCCAGCAGTTCCAGAGCAGAACGATGTACATAGG | 1167 |
| Db | 3382 | CGGCAGGACCTATTGGAGGCGACTTTCCAGCAGTTCCAGAGCAGAACGATGTACATAGG | 3441 |
| Qy | 1168 | GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGGAGACTGTA | 1227 |
| Db | 3442 | GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGGAGACTGTA | 3501 |
| Qy | 1228 | CGAATATTTCTGACAGAGCAGCCTTTTGAAGGACTAGAGAAAACTCTACAGGAGCCGAGA | 1287 |
| Db | 3502 | CGAATATTTCTGACAGAGCAGCCTTTTGAAGGACTAGAGAAAACTCTACAGGAGCCGAGA | 3561 |
| Qy | 1288 | GAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG | 1347 |
| Db | 3562 | GAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG | 3621 |
| Qy | 1348 | GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCGCTGACTGGCAGAGAAAAATA | 1407 |
| Db | 3622 | GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCGCTGACTGGCAGAGAAAAATA | 3681 |
| Qy | 1408 | GATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGSCCAGGATGAGCTGGACCTCAAG | 1467 |
| Db | 3682 | GATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGSCCAGGATGAGCTGGACCTCAAG | 3741 |
| Qy | 1468 | CTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGAC | 1527 |
| Db | 3742 | CTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGAC | 3801 |
| Qy | 1528 | TCTCTCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAAAATTGGCCCTCTGAAA | 1587 |
| Db | 3802 | TCTCTCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAAAATTGGCCCTCTGAAA | 3861 |
| Qy | 1588 | GAGAACCTGAGCCACGTCAATGACCTTGTCTCGCCAGCTTACCACCTTGGGCAATCAGCTC | 1647 |
| Db | 3862 | GAGAACCTGAGCCACGTCAATGACCTTGTCTCGCCAGCTTACCACCTTGGGCAATCAGCTC | 3921 |
| Qy | 1648 | TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAGGTG | 1707 |
| Db | 3922 | TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAGGTG | 3981 |
| Qy | 1708 | GCCGTCGAGGACCGAGTCAGGAGCTGCATGAAGCCCAAGGACCTTGTGTCAGCATCT | 1767 |
| Db | 3982 | GCCGTCGAGGACCGAGTCAGGAGCTGCATGAAGCCCAAGGACCTTGTGTCAGCATCT | 4041 |
| Qy | 1768 | CAGCACTTCTTTCCACGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAAAAA | 1827 |
| Db | 4042 | CAGCACTTCTTTCCACGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAAAAA | 4101 |
| Qy | 1828 | GTGCCCTACTATATCAACCCAGAGACTCAAAACAACCTTGTGGGACCATCCCAAATGACA | 1887 |
| Db | 4102 | GTGCCCTACTATATCAACCCAGAGACTCAAAACAACCTTGTGGGACCATCCCAAATGACA | 4161 |
| Qy | 1888 | GAGCTCTACCAAGTCTTTAGCTGACCTGAATAATGTGACATCTCAGCTTATAGGACTGCC | 1947 |
| Db | 4162 | GAGCTCTACCAAGTCTTTAGCTGACCTGAATAATGTGACATCTCAGCTTATAGGACTGCC | 4221 |
| Qy | 1948 | ATGAAACTCCGAAGACTGCAAGAGGCCCTTTTGTGATCTCTTGAGCCTGTGAGCTGCA | 2007 |

| | | | |
|----------------------------|------|--|------|
| Db | 4222 | ATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTTCAGCTGCA | 4281 |
| QY | 2008 | TGTGATGCCTTGGACCAAGCACAACCTCAAGCAAAATGACCAGCCCATGGATATCCTTGCA | 2067 |
| Db | 4282 | TGTGATGCCTTGGACCAAGCACAACCTCAAGCAAAATGACCAGCCCATGGATATCCTTGCA | 4341 |
| QY | 2068 | ATTATTAAATTGTTGACCACTATTTATGACCGCCTGGAGCAAGAGCACAACAATTGTC | 2127 |
| Db | 4342 | ATTATTAAATTGTTGACCACTATTTATGACCGCCTGGAGCAAGAGCACAACAATTGTC | 4401 |
| QY | 2128 | AACGTCCCTCTCTGCGTGGATATGTGCTGAACTGGCTGCTGTAATGTTTATGATACGGGA | 2187 |
| Db | 4402 | AACGTCCCTCTCTGCGTGGATATGTGCTGAACTGGCTGCTGTAATGTTTATGATACGGGA | 4461 |
| QY | 2188 | CGAACAGGAGGATCCGTGTCCTGTCCTTTTAAAACTGGCATCAATTCCTGTGTAAAGCA | 2247 |
| Db | 4462 | CGAACAGGAGGATCCGTGTCCTGTCCTTTTAAAACTGGCATCAATTCCTGTGTAAAGCA | 4521 |
| QY | 2248 | CATTGGAAGACAAAGTACAGATACCTTTTCAAGCAAAGTGGCAAAGTTCAACAGGATTTTGT | 2307 |
| Db | 4522 | CATTGGAAGACAAAGTACAGATACCTTTTCAAGCAAAGTGGCAAAGTTCAACAGGATTTTGT | 4581 |
| QY | 2308 | GACCAGCGCAGGTGGGCCTCCTTCTGCAATGATCTATCCAAATTCCAAAGACAGTTGGST | 2367 |
| Db | 4582 | GACCAGCGCAGGTGGGCCTCCTTCTGCAATGATCTATCCAAATTCCAAAGACAGTTGGST | 4641 |
| QY | 2368 | GAAGTTGCATCCTTTGGGGGCAGTAACATTGAGC | 2401 |
| Db | 4642 | GAAGTTGCATCCTTTGGGGGCAGTAACATTGAGC | 4675 |
| RESULT 3 | | | |
| AX114289 | | | |
| LOCUS | | | |
| DEFINITION | | | |
| ACCESSION | | | |
| VERSION | | | |
| KEYWORDS | | | |
| SOURCE | | | |
| ORGANISM | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| FEATURES | | | |
| source | | | |
| misc_feature | | | |
| misc_feature | | | |
| BASE COUNT | | | |
| ORIGIN | | | |
| Query Match | | | |
| Best Local Similarity | | | |
| Matches 1750; Conservative | | | |
| QY | 628 | TGGGCAACATCTGTAGATGGACAGAACCCGCTGGTCTTTTACAAGACACTCATAGA | 687 |
| Db | 2902 | TGGAGAAGCATTCATAAAAGGGTGTAGTGACGAGAGGCTGCTTTGGAAGAACTCATAGA | 2961 |
| QY | 688 | TTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGGCTGGCTTACAGAAGCTGAA | 747 |
| Db | 2962 | TTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGGCTGGCTTACAGAAGCTGAA | 3021 |

| | | | |
|----|------|---|------|
| QY | 748 | ACAACTGCCAATGTCTCTACAGGATGTCTACCCGTAAAGAAAGSTCTCTAGAACTCCAAG | 807 |
| Db | 3022 | ACAACTGCCAATGTCTCTACAGGATGTCTACCCGTAAAGAAAGSTCTCTAGAACTCCAAG | 3081 |
| QY | 808 | GGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACA | 867 |
| Db | 3082 | GGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACA | 3141 |
| QY | 868 | GATGTTTATCAACAACCTGGATGAAACACAGCCAAAAATCCTGAGATCCCTGGAAGTTCC | 927 |
| Db | 3142 | GATGTTTATCAACAACCTGGATGAAACACAGCCAAAAATCCTGAGATCCCTGGAAGTTCC | 3201 |
| QY | 928 | GATGATGCAGTCTCTTTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAGTGAACCTT | 987 |
| Db | 3202 | GATGATGCAGTCTCTTTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAGTGAACCTT | 3261 |
| QY | 988 | CGGAAAAAGTCTCTCAACATTAGGTCCCATTGGAAGCCAGTTTGCACAGTGGAAGCGT | 1047 |
| Db | 3262 | CGGAAAAAGTCTCTCAACATTAGGTCCCATTGGAAGCCAGTTTGCACAGTGGAAGCGT | 3321 |
| QY | 1048 | CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGTGAAGATGATGAATTAAGC | 1107 |
| Db | 3322 | CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGTGAAGATGATGAATTAAGC | 3381 |
| QY | 1108 | CGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGG | 1167 |
| Db | 3382 | CGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGG | 3441 |
| QY | 1168 | GCCTTCAAGAGGGGAATTGAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTA | 1227 |
| Db | 3442 | GCCTTCAAGAGGGGAATTGAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTA | 3501 |
| QY | 1228 | CGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACTTACCAGGAGCCCAGA | 1287 |
| Db | 3502 | CGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACTTACCAGGAGCCCAGA | 3561 |
| QY | 1288 | GAGCTGCCTCCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAG | 1347 |
| Db | 3562 | GAGCTGCCTCCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAG | 3621 |
| QY | 1348 | GAGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATA | 1407 |
| Db | 3622 | GAGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATA | 3681 |
| QY | 1408 | GATGAGACCTTGAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAG | 1467 |
| Db | 3682 | GATGAGACCTTGAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAG | 3741 |
| QY | 1468 | CTGCGCAAGCTGAGGTGATCAAGGATCTCTGGCAGCCCGTGGCGATCTCCTCATTTGAC | 1527 |
| Db | 3742 | CTGCGCAAGCTGAGGTGATCAAGGATCTCTGGCAGCCCGTGGCGATCTCCTCATTTGAC | 3801 |
| QY | 1528 | TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAA | 1587 |
| Db | 3802 | TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAA | 3861 |
| QY | 1588 | GAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACTTTGGGCATTCAGCTC | 1647 |
| Db | 3862 | GAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACTTTGGGCATTCAGCTC | 3921 |
| QY | 1648 | TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAAGTGAAGCTTCTGCAGGTG | 1707 |
| Db | 3922 | TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAAGTGAAGCTTCTGCAGGTG | 3981 |
| QY | 1708 | GCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGACTTTGGTCCAGCATCT | 1767 |
| Db | 3982 | GCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGACTTTGGTCCAGCATCT | 4041 |
| QY | 1768 | CAGCACTTTCTTTCCACGTCTGTCCAGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA | 1827 |
| Db | 4042 | CAGCACTTTCTTTCCACGTCTGTCCAGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA | 4101 |

```
QY 1828 GTGCCCTACTATATCAACACGAGACTCAAAACAATTGCTGGGACCATCCCAAAATGACA 1887
|
|
|
Db 4102 GTGCCCTACTATATCAACACGAGACTCAAAACAATTGCTGGGACCATCCCAAAATGACA 4161
|
|
|
QY 1888 GAGCTCTACCAAGTCTTTAGCTGACCTGAATATGTTCAGATTCTCAGCTTATAGGACTGCC 1947
|
|
|
Db 4162 GAGCTCTACCAAGTCTTTAGCTGACCTGAATATGTTCAGATTCTCAGCTTATAGGACTGCC 4221
|
|
|
QY 1948 ATGAAACTCCGAAGACTGACAGAGGCCCCCTTTGCTGGATCTCTTGAGCTGTTCAGCTGCA 2007
|
|
|
Db 4222 ATGAAACTCCGAAGACTGACAGAGGCCCCCTTTGCTGGATCTCTTGAGCTGTTCAGCTGCA 4281
|
|
|
QY 2008 TGTGATGCCCTTGGACAGCACAACTCAAGCAAAATGACAGCCCCATGGATATCCTGCAG 2067
|
|
|
Db 4282 TGTGATGCCCTTGGACAGCACAACTCAAGCAAAATGACAGCCCCATGGATATCCTGCAG 4341
|
|
|
QY 2068 ATTATTAATTGTTTGACCACTATTATGACCGCTGGAGCAAGAGCACAACTTTGGTC 2127
|
|
|
Db 4342 ATTATTAATTGTTTGACCACTATTATGACCGCTGGAGCAAGAGCACAACTTTGGTC 4401
|
|
|
QY 2128 AACGTCCCTCTCTGCTGGATATGTGTCTGAAGTGGCTGCTGAATGTTTATGATACGGGA 2187
|
|
|
Db 4402 AACGTCCCTCTCTGCTGGATATGTGTCTGAAGTGGCTGCTGAATGTTTATGATACGGGA 4461
|
|
|
QY 2188 CGAACAGGGAGGATCCGTGTCCTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 2247
|
|
|
Db 4462 CGAACAGGGAGGATCCGTGTCCTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 4521
|
|
|
QY 2248 CATTTGGAAGACAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGT 2307
|
|
|
Db 4522 CATTTGGAAGACAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGT 4581
|
|
|
QY 2308 GACCAGCGCAGGCTGGGCTCCTTCTGATGATTTCTATCCAAATTCGAAGACAGTTGGGT 2367
|
|
|
Db 4582 GACCAGCGCAGGCTGGGCTCCTTCTGATGATTTCTATCCAAATTCGAAGACAGTTGGGT 4641
|
|
|
QY 2368 GAAAGTTGCATCCTTTGGGGCAGTAACATTGAGC 2401
|
|
|
Db 4642 GAAAGTTGCATCCTTTGGGGCAGTAACATTGAGC 4675
|
|
|
```

```
RESULT 4
AX538622
LOCUS AX538622 8689 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 42 from Patent WO0229056.
ACCESSION AX538622
VERSION AX538622.1 GI:25271171
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
```

```
REFERENCE
1
AUTHORS Chamberlain,J.S. and Harper,S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 42 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
source
1. .8689
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"
BASE COUNT 2721 a 1804 c 1861 g 2303 t
ORIGIN
```

```
Query Match 72.3%; Score 1735.6; DB 6; Length 8689;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 628 TGGGCAACATCTGTAGATGGACAGAGACCGCTGGTCTTTTACAAGACACTCATAGA 687
|
|
|
Db 2948 TGGAGAAGCATTCATAAAAGGGTGAGTGACCGAGAGGCTGCTTTGGAAGAACTCATAGA 3007
|
|
|
```

```
QY 688 TTAATGCAACAGTTCCCTCGACCTGGAAAGTTTCTTGCTGGCTTACAGAAGCTGAA 747
|
|
|
Db 3008 TTAATGCAACAGTTCCCTCGACCTGGAAAGTTTCTTGCTGGCTTACAGAAGCTGAA 3067
|
|
|
QY 748 ACAACTGCCAATGTCTACAGGATGTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAG 807
|
|
|
Db 3068 ACAACTGCCAATGTCTACAGGATGTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAG 3127
|
|
|
QY 808 GGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACA 867
|
|
|
Db 3128 GGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACA 3187
|
|
|
QY 868 GATGTTTATCACAACTGGATGAAACAGCAAGCAAAATCCTGAGATCCCTGGAAGTTCC 927
|
|
|
Db 3188 GATGTTTATCACAACTGGATGAAACAGCAAGCAAAATCCTGAGATCCCTGGAAGTTCC 3247
|
|
|
QY 928 GATGATGCAGTCTCTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTT 987
|
|
|
Db 3248 GATGATGCAGTCTCTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTT 3307
|
|
|
QY 988 CGGAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAAGTGAAGCGT 1047
|
|
|
Db 3308 CGGAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAAGTGAAGCGT 3367
|
|
|
QY 1048 CTGCACCTTTCTCTCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 1107
|
|
|
Db 3368 CTGCACCTTTCTCTCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 3427
|
|
|
QY 1108 CGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAAAGCAGATGTACATAGG 1167
|
|
|
Db 3428 CGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAAAGCAGATGTACATAGG 3487
|
|
|
QY 1168 GCCTTCAAGAGGGAATTGAAAACCTAAAGAACCTGTAATCATGATGACTCTTGGAGACTGTA 1227
|
|
|
Db 3488 GCCTTCAAGAGGGAATTGAAAACCTAAAGAACCTGTAATCATGATGACTCTTGGAGACTGTA 3547
|
|
|
QY 1228 CGAATATTTCTGACAGAGCAGCCTTTTGGAGGACTAGAGAAACCTTACAGAGGCCAGA 1287
|
|
|
Db 3548 CGAATATTTCTGACAGAGCAGCCTTTTGGAGGACTAGAGAAACCTTACAGAGGCCAGA 3607
|
|
|
QY 1288 GAGCTGCCCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 1347
|
|
|
Db 3608 GAGCTGCCCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 3667
|
|
|
QY 1348 GAGGTCAATGAGTGGGAAAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 1407
|
|
|
Db 3668 GAGGTCAATGAGTGGGAAAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 3727
|
|
|
QY 1408 GATGAGACCTTGAAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAG 1467
|
|
|
Db 3728 GATGAGACCTTGAAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAG 3787
|
|
|
QY 1468 CTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCCGCTGGGCGATCTCCTCATTTGAC 1527
|
|
|
Db 3788 CTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCCGCTGGGCGATCTCCTCATTTGAC 3847
|
|
|
QY 1528 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTTGGCCTCTGAAA 1587
|
|
|
Db 3848 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTTGGCCTCTGAAA 3907
|
|
|
QY 1588 GAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCATTGGGCAATTCAGCTC 1647
|
|
|
Db 3908 GAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCATTGGGCAATTCAGCTC 3967
|
|
|
QY 1648 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTG 1707
|
|
|
Db 3968 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTG 4027
|
|
|
QY 1708 GCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCCCACAGGACTTTGGTCCAGCATCT 1767
|
|
|
Db 4028 GCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCCCACAGGACTTTGGTCCAGCATCT 4087
|
|
|
QY 1768 CAGCACCTTTCTTTCCAGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 1827
|
|
|
```


Db 4088 CAGCACTTTCTTCCACGCTGTGTCAGGGTCCCTGGGAGAGACCATCTCGCCAAACAAA 4147
QY 1828 GTGCCCCACTACTATATCAACCACGAGACTCAAAACAACTTGCTGGGACCATCCCAAATGACA 1887
Db 4148 GTGCCCCACTACTATATCAACCACGAGACTCAAAACAACTTGCTGGGACCATCCCAAATGACA 4207
QY 1888 GAGCTCTACCACTCTTTAGCTGACCTGAATAATGTGATGATCTCAGCTTATAGGACTGCC 1947
Db 4208 GAGCTCTACCACTCTTTAGCTGACCTGAATAATGTGATGATCTCAGCTTATAGGACTGCC 4267
QY 1948 ATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCA 2007
Db 4268 ATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCA 4327
QY 2008 TGTGATGCTTGGACCAAGCACAACTCAAGCAAAATGACAGCCCATGGATATCCTGCGAG 2067
Db 4328 TGTGATGCTTGGACCAAGCACAACTCAAGCAAAATGACAGCCCATGGATATCCTGCGAG 4387
QY 2068 ATTATTAATTGTTGACCACTATTTATGACCGCCTGGAGCAAGAGCACAAATTTGGTC 2127
Db 4388 ATTATTAATTGTTGACCACTATTTATGACCGCCTGGAGCAAGAGCACAAATTTGGTC 4447
QY 2128 AACGTCCCTCTCTGCGTGGATATGTGTGAACCTGGCTGCTGAATGTTTATGATACGGGA 2187
Db 4448 AACGTCCCTCTCTGCGTGGATATGTGTGAACCTGGCTGCTGAATGTTTATGATACGGGA 4507
QY 2188 CGAACAGGGAGGATCCGTGTCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAGCA 2247
Db 4508 CGAACAGGGAGGATCCGTGTCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAGCA 4567
QY 2248 CATTTGGAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGT 2307
Db 4568 CATTTGGAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGT 4627
QY 2308 GACCAGCGCAGGCTGGGCCTCCTTCTGATGATTTATCCAAATTTCAAGACAGTTGGGT 2367
Db 4628 GACCAGCGCAGGCTGGGCCTCCTTCTGATGATTTATCCAAATTTCAAGACAGTTGGGT 4687
QY 2368 GAAATTGCATCCTTTTGGGGGCGAGTAACATTGAGC 2401
Db 4688 GAAATTGCATCCTTTTGGGGGCGAGTAACATTGAGC 4721

RESULT 5
AX538624
LOCUS AX538624 11443 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 44 from Patent WO0229056.
ACCESSION AX538624
VERSION AX538624.1 GI:25271175
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Chamberlain,J.S. and Harper,S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 44 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
source
1. .11443
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"
BASE COUNT 3707 a 2339 c 2502 g 2895 t
ORIGIN
Query Match 72.3%; Score 1735.6; DB 6; Length 11443;
Best Local Similarity 98.6%; Pred.No. 0;
Matches 1750; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 628 TGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACACTCATAGA 687

Db 5702 TGGAGAAGCATTCATAAAAGGTGAGTGAGCGAGAGGCTGCTTTGGAAGAACTCATAGA 5761
QY 688 TTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGGCTGGCTTACAGAAGCTGAA 747
Db 5762 TTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGGCTGGCTTACAGAAGCTGAA 5821
QY 748 ACAACTGCCAATGTCTTACAGGATGCTACCCGTAAAGGAAAGGCTCCTAGAAGACTCCAAG 807
Db 5822 ACAACTGCCAATGTCTTACAGGATGCTACCCGTAAAGGAAAGGCTCCTAGAAGACTCCAAG 5881
QY 808 GGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACA 867
Db 5882 GGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACA 5941
QY 868 GATGTTTATCACAACTTGGATGAAACAGCCAAAATCCTGAGATCCCTGGAAGGTTCC 927
Db 5942 GATGTTTATCACAACTTGGATGAAACAGCCAAAATCCTGAGATCCCTGGAAGGTTCC 6001
QY 928 GATGATGAGTCTCTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAACCT 987
Db 6002 GATGATGAGTCTCTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAACCT 6061
QY 988 CGGAAAAAGTCTCTCAACATTAGTCCCATTTTGGAAAGCCAGTTCTGACCAGTGGAAAGCCT 1047
Db 6062 CGGAAAAAGTCTCTCAACATTAGTCCCATTTTGGAAAGCCAGTTCTGACCAGTGGAAAGCCT 6121
QY 1048 CTGCACCTTTCTTGCAGGAACCTTCTGGTGTGGCTACAGTGAAGATGATGAATTAAGC 1107
Db 6122 CTGCACCTTTCTTGCAGGAACCTTCTGGTGTGGCTACAGTGAAGATGATGAATTAAGC 6181
QY 1108 CGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTTCAGAAAGCAGAACGATGTACATAGG 1167
Db 6182 CGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTTCAGAAAGCAGAACGATGTACATAGG 6241
QY 1168 GCCTTCAGAGGGGAATTGAAAACCTTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTA 1227
Db 6242 GCCTTCAGAGGGGAATTGAAAACCTTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTA 6301
QY 1228 CGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGAGGCCCAGA 1287
Db 6302 CGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGAGGCCCAGA 6361
QY 1288 GAGCTGCCTCTGAGGAGAGAGCCCAAGATGTCACTCGCTTCTACGAAAGCAGGCTGAG 1347
Db 6362 GAGCTGCCTCTGAGGAGAGAGCCCAAGATGTCACTCGCTTCTACGAAAGCAGGCTGAG 6421
QY 1348 GAGGTCAATACTAGTGGGAAAAATTGAACCTGCACCTCGCTGACTGGCAGAGAAAAATA 1407
Db 6422 GAGGTCAATACTAGTGGGAAAAATTGAACCTGCACCTCGCTGACTGGCAGAGAAAAATA 6481
QY 1408 GATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAG 1467
Db 6482 GATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAG 6541
QY 1468 CTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGAC 1527
Db 6542 CTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGAC 6601
QY 1528 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCAGAGAGAAATTTGCGCCTCTGAAA 1587
Db 6602 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCAGAGAGAAATTTGCGCCTCTGAAA 6661
QY 1588 GAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTCAGCTC 1647
Db 6662 GAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTCAGCTC 6721
QY 1648 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTG 1707
Db 6722 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTG 6781
QY 1708 GCCGTGAGGACCGGATCAGGACAGCTGATGAAGCCCAAGGACCTTTGGTCCAGCATCT 1767

Db 6782 GCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCT 6841
QY 1768 CAGCACTTTCTTTCCACGTCCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 1827
Db 6842 CAGCACTTTCTTTCCACGTCCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 6901
QY 1828 GTGCCCTACTATATCAACCACGAGACTCAAAACAACCTTGCTGGGACCATCCAAAATGACA 1887
Db 6902 GTGCCCTACTATATCAACCACGAGACTCAAAACAACCTTGCTGGGACCATCCAAAATGACA 6961
QY 1888 GAGCTCTACCAAGTCTTTAGCTGACCTGAAATAATGTTCAGATTCTCAGCTTATAGGACTGCC 1947
Db 6962 GAGCTCTACCAAGTCTTTAGCTGACCTGAAATAATGTTCAGATTCTCAGCTTATAGGACTGCC 7021
QY 1948 ATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTTCAGCTGCA 2007
Db 7022 ATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTTCAGCTGCA 7081
QY 2008 TGTGATGCCCTTGGACCAAGCAACAACCTCAAGCAAAAATGACCGCCCATGGATATCCTTGCGAG 2067
Db 7082 TGTGATGCCCTTGGACCAAGCAACAACCTCAAGCAAAAATGACCGCCCATGGATATCCTTGCGAG 7141
QY 2068 ATTATTAATTGTTGACCACCTATTTATGACCGCCTGGAGCAAGACGACACAACAATTTGGTC 2127
Db 7142 ATTATTAATTGTTGACCACCTATTTATGACCGCCTGGAGCAAGACGACACAACAATTTGGTC 7201
QY 2128 AACGTCCCTCTCTGCGTGGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGA 2187
Db 7202 AACGTCCCTCTCTGCGTGGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGA 7261
QY 2188 CGAACAGGGAGGATCCGTGTCTCTTTTAAAACTGGCATCATATTTCCCTGTGTAAAGCA 2247
Db 7262 CGAACAGGGAGGATCCGTGTCTCTTTTAAAACTGGCATCATATTTCCCTGTGTAAAGCA 7321
QY 2248 CATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGT 2307
Db 7322 CATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGT 7381
QY 2308 GACCAGCGCAGCTGGGCCTCCTTCTGTCATGATTTCTATCCAAATTTCAAGACAGTTGGGT 2367
Db 7382 GACCAGCGCAGCTGGGCCTCCTTCTGTCATGATTTCTATCCAAATTTCAAGACAGTTGGGT 7441
QY 2368 GAAGTTGCATCCTTTGGGGGCAGTAAACATTGAGC 2401
Db 7442 GAAGTTGCATCCTTTGGGGGCAGTAAACATTGAGC 7475

RESULT 6
AX538627
LOCUS AX538627 12057 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 47 from Patent WO0229056.
ACCESSION AX538627
VERSION AX538627.1 GI:25271181
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE
1
AUTHORS Chamberlain,J.S. and Harper,S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 47 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
source
1. .12057
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"
BASE COUNT 4020 a 2456 c 2790 g 2791 t
ORIGIN

Query Match 72.3%; Score 1735.6; DB 6; Length 12057;
Best Local Similarity 98.6%; Pred. No. 0;

Matches 1750; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 628 TGGGCAAAACATCTGTAGATGGACAGAAAGACCGCTGGGTTCTTTTACAAGACACTCATAGA 687
Db 8216 TGGAGAAAGCATTCATAAAAGGGTGAGTACGCGAGAGGCTGCTTTTGGAAAGAACTCATAGA 8275
QY 688 TTAAGTCAACAGTTCCCTGGACCTTGAAAGGTTCTTGGCTTACAGAAAGCTGAA 747
Db 8276 TTAAGTCAACAGTTCCCTGGACCTTGAAAGGTTCTTGGCTTACAGAAAGCTGAA 8335
QY 748 ACAAGTCCCAATGCTTACAGGATGCTACCCGTAAGGAAAGGCTCTTGAAGACTCCAAG 807
Db 8336 ACAAGTCCCAATGCTTACAGGATGCTACCCGTAAGGAAAGGCTCTTGAAGACTCCAAG 8395
QY 808 GGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAATTTGAAGCTCACACA 867
Db 8396 GGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAATTTGAAGCTCACACA 8455
QY 868 GATGTTTATCAACAACCTGGATGAAACAGCCCAAAAATCCTGAGATCCCTGGAAGGTTCC 927
Db 8456 GATGTTTATCAACAACCTGGATGAAACAGCCCAAAAATCCTGAGATCCCTGGAAGGTTCC 8515
QY 928 GATGATGCAGTCTGTTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCT 987
Db 8516 GATGATGCAGTCTGTTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCT 8575
QY 988 CGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCGT 1047
Db 8576 CGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCGT 8635
QY 1048 CTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 1107
Db 8636 CTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 8695
QY 1108 CGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTTCCAGAACGATGATGATGATGATGATG 1167
Db 8696 CGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTTCCAGAACGATGATGATGATGATGATG 8755
QY 1168 GCCTTCAAGAGGGAATTGAAACCTTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTA 1227
Db 8756 GCCTTCAAGAGGGAATTGAAACCTTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTA 8815
QY 1228 CGAATATTTCTGACAGAGCAGCCTTTTGGAAAGGACTAGAGAACTCTACCAGGAGCCAGA 1287
Db 8816 CGAATATTTCTGACAGAGCAGCCTTTTGGAAAGGACTAGAGAACTCTACCAGGAGCCAGA 8875
QY 1288 GAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 1347
Db 8876 GAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 8935
QY 1348 GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 1407
Db 8936 GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 8995
QY 1408 GATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAG 1467
Db 8996 GATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAG 9055
QY 1468 CTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGAC 1527
Db 9056 CTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGAC 9115
QY 1528 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAA 1587
Db 9116 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAA 9175
QY 1588 GAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTCAGCTC 1647
Db 9176 GAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTCAGCTC 9235
QY 1648 TCACCGTATAACCTCAGCAGCTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTG 1707
Db 9236 TCACCGTATAACCTCAGCAGCTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTG 9295

QY 1708 GCCGTCGAGACCGAGTCAGGCAGCTGCATGAAGCCACACAGGACATTGGTCACGATCT 1767
|||||
Db 9296 GCCGTCGAGACCGAGTCAGGCAGCTGCATGAAGCCACACAGGACATTGGTCACGATCT 9355
|||||
QY 1768 CAGCACTTTCTTCCACGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 1827
|||||
Db 9356 CAGCACTTTCTTCCACGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 9415
|||||
QY 1828 GTGCCCTACTATATCAACACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAATGACA 1887
|||||
Db 9416 GTGCCCTACTATATCAACACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAATGACA 9475
|||||
QY 1888 GAGCTCTACAGCTCTTTAGCTGACCTGAATAATGTTCAGATTTCTCAGCTTATAGGACTGCC 1947
|||||
Db 9476 GAGCTCTACAGCTCTTTAGCTGACCTGAATAATGTTCAGATTTCTCAGCTTATAGGACTGCC 9535
|||||
QY 1948 ATGAAACTCCGAAGACTGCAGAGGCGCCCTTGCTTGGATCTCTTGAGCCTGTACGCTGCA 2007
|||||
Db 9536 ATGAAACTCCGAAGACTGCAGAGGCGCCCTTGCTTGGATCTCTTGAGCCTGTACGCTGCA 9595
|||||
QY 2008 TGTGATGCTTGGACACGAGCACAACCTCAAGCAAAATGACACAGCCCATGGATATCTCTGCAG 2067
|||||
Db 9596 TGTGATGCTTGGACACGAGCACAACCTCAAGCAAAATGACACAGCCCATGGATATCTCTGCAG 9655
|||||
QY 2068 ATTATTAATTGTTTGACCACTATTATTATGACCGCCTGGAGCAAGAGCACAACAATTGGTC 2127
|||||
Db 9656 ATTATTAATTGTTTGACCACTATTATTATGACCGCCTGGAGCAAGAGCACAACAATTGGTC 9715
|||||
QY 2128 AACGTCCTCTCTGCGTGGATATGTCTGAAGCTGGCTGGTGAATGTTTATGATACGGGA 2187
|||||
Db 9716 AACGTCCTCTCTGCGTGGATATGTCTGAAGCTGGCTGGTGAATGTTTATGATACGGGA 9775
|||||
QY 2188 CGAACAGGAGGATCCGTGCTCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 2247
|||||
Db 9776 CGAACAGGAGGATCCGTGCTCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 9835
|||||
QY 2248 CATTTGGAACACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGT 2307
|||||
Db 9836 CATTTGGAACACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGT 9895
|||||
QY 2308 GACCAGCGCAGGCTGGCCTCCTCTGTCATGATTCTATCCAAATTCACAGACAGTTGGGT 2367
|||||
Db 9896 GACCAGCGCAGGCTGGCCTCCTCTGTCATGATTCTATCCAAATTCACAGACAGTTGGGT 9955
|||||
QY 2368 GAAGTTGCATCCTTTGGGGGCAGTAACATTGAGC 2401
|||||
Db 9956 GAAGTTGCATCCTTTGGGGGCAGTAACATTGAGC 9989
|||||

RESULT 7
HSDMDR HSDMDR 12446 bp mRNA linear PRI 12-SEP-1993
LOCUS Human mRNA for dystrophin.
DEFINITION X14298
ACCESSION
VERSION X14298.1 GI:30845
KEYWORDS Dmd gene; Duchenne muscular dystrophy; dystrophin.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 12446)
AUTHORS Rosenthal,A., Speer,A., Billwitz,H., Cross,G.S., Forrest,S.M. and Davies,K.E.
TITLE Two human cDNA molecules coding for the Duchenne muscular dystrophy (DMD) locus are highly homologous
JOURNAL Nucleic Acids Res. 17 (13), 5391 (1989)
MEDLINE 89345106
PUBMED 2668885
REFERENCE 2 (bases 1 to 12446)
AUTHORS Rosenthal,A.
TITLE Direct Submission
JOURNAL Submitted (09-FEB-1989) Rosenthal A., Akademie der Wissenschaften

der DDR, Zentralinstitut fuer Molekularbiologie, Robert-Roessle Str.10, 1115 Berlin Buch, DDR
see also M18533 and M20250 for Dmd segs.; discrepancies compared to M18533 cDNA were located at x14298 pos. 496, 1772, 1965, 2449, 3687, 4229, 4504, 5075, 5332, 5630 and 7194.
FEATURES
Location/Qualifiers
1..12446
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X chromosomal, Xp21."
/tissue_type="muscle"
/dev_stage="fetal and adult."
99..11156
/note="dystrophin (AA 1 - 3685)"
/codon_start=1
/protein_id="CAA32479.1"
/db_xref="GI:30846"
/db_xref="SWISS-PROT:P11532"
/translation="MLWEEVEDCYEREDVQKKTFTKWNNAQFSKFGKHIEHLFSDL QDGRLLDLLLEGLTGQKLPKEGSTRVHALNNVNKALRVLQNNVDLVNIGSTDIVDG NHKLTGLIWNIIILHWQVKNVMKNIMAGLQPTNSEKILLSWVRQSRFNYPQVNVINF TSWSDGLALNALITHSRPDLFDWNSVVCOSATQRLHAFNARIYQLGIEKLLDPEDV DITYPDKKSILMYITSLFQVLPQQVSIETAIQEVMLPRPPKVTKEEHFQLHHQMHYSQ QITVSLAQGYERTSPKPRFKSYAYTAAAYTSDTPSPFPSSHLEAPEDKSFSSSL MESEVNLDRYQTALBEVLSWLSAEDTLQAQGEISNDVEVVKDQFHHHEGYMDLTAH OGRVGNILQIGSLITGKLTSEDETEVEQMNLLNSRWECLRVASMEKQSNLHRVLM DLQNKLKELNDWLTKEERTKMEEPGLDLEDLKRQVQKHVLQEDLEQEQRVFN SLTHMVVVVDESSGDHATAALEEQLVGDRWANI CRWTEDRWVLLQDILLKWQRLTE EQCLFSAWLSEKEDAVNKIHTTGFKDQDNEMSSLQKLAVLKADLEKKQSMGKLYSIK ODLLSTLKNKSVTQKTEAWLDNFARCDNLVQKLEKSTAQISOAVTTTQPSLTQTITVM ETVTVTTRTEQILVKHAQEEELPPPPPKKQITVDSEIRKRLDVIDTELHSWITRSEA VLQSPFAIFRKEGNFSDLKEKVNAIEREAKFKRLQDASRSQALVEQMVNEGVNA DSIKQASEQLNSRWIEFCQLLSERLNWLEYQNNIIAFYNQLQQLQEQMTTAEANWLKIQ PTPPSEPTAIKSQLKICKDEVNRLSGLQPOIERLKIQSIALKEKGQGMFLDADFVAF TNHFQKQFSDVQAREKELQTI FDTLPPMYQETMSAIRTWVQQSETKLSIPQLSVTDY EIMEORLGELOALQSSLQEQSGLYLSTTVTKEMSKAPSEISRKYQSEFEIEGRWK KLSQLVEHCQKLEEQMNKLRKIQNHIIQTLKKWMAEVDVFLKEWPALGDSEILKKQL XQCRLLVSDIOTIQPSLNSVNEGQKIKNAEPEPASRLETETELKELNTQWDHMCQQVY ARKEALKGLEKTVSLQKDLSEMHWMTOAEIEYLDERDFEYKTPDELQKAFPEMKRAK BEAQOKEAKVLLTESVNSVIAQAPFVAQEALKKELETLTNTYNQWLCTRLNGKCKTLE EWACWHELLSYLEKANKWLNEVEFKLTETENIPGGAEEISEVLSLENLMRHSNPN NQIRILAQTLTDGGVMDLINEELETENFNRWRELHEEAARRQKLEQSIQSAQETENS LHLIQESLTFIDKOLAAVIAADKVDAAQMPQOAKIQSDLTSHIEISLEEMKGNQKGEA AQRVLSQIDVAQKKLDQVSMKFRFLQKPANFEQRLQESKMILDEVKMLPALETKSVE QEVVQSQLNHCNLYKLSLSEVSEVMVITGRQIVQKQOTENPKELDERVTALKLHY NELGAKVTERKQQLKCLSRKMRKEMNVLTETLAATDMELTKRSAVEGMPNSLDSE VAWGKATQKEIEKQVHLKSI TEVGEALXTVLGKETLVLEDKLSLLNSNWI AVTSRAE EWLNLLEYQKHEMTFDQNVVDHITKWI IQADTLTLDSEKKKPPQKEDVLKRLKAE LND IRPKVDS TRDQAANLMAHGDHCRKLVEPOISELNHFAAISHRIKTKASIP LKEL E QNSDIOKLLPLEAEITQQGVNLKEEDFNKMDNEDNEGTVKELQRLQDNLQORITDER KSEEEKIKQQLLTQKHNALKDLRSORRKKAL EISHQWYQYKQADDLKCLDDIEKKL ASLPEPRDERKIEIDRELQKKEELNAVRRQAEGLSEYSPSTYLTETHTVSQALLVEEQ IESKFAQFRRLNFAQIHTVREETMMVMTEDMPLEISIVSPSTYLTETHTVSQALLVEEQ LLNAPDLCAKQFEDLFKQEESLKNIKDSLQOSSGRIDI IHSKKTAAALQSATPVERVKL QEALSQLDFQWKVNMYKDRQGRFDRSVEKWRFRHYDIKIFNQWLTEAEQFLRKTQI PENWEHAKYKWLKBLQDGIGQRTVVRTLNATGEEIIQSSKTASILQEKLGSLNL RWQEVCKQLSDRKKRLEEQNILSEFQRLNEFVWLWEEADNIASIPLEPKEQQLKE KLEQVKLLVEELPLRQILKQLNETGGPVLVSAPISPEEQKLENKLKQTNLQWIKVS RALPEKQGEIEAQIKDLGQLEKKLEDLEEQLNHLHLWLSPIRNQLEIYNQPNQEGPFD VKETEIAVQAKQPDVBEILSKGHLYKEKPATQPVKRLLEDLSSEWKA VNRLLQELRA KQPD LAPGLTTIGASQTQTVTLVTPVTKETAISKLEMPSSLMLEVPALADFNRAWT ELTDWLSLLDQVIKSORVMVGDL EDINEMI IKQKATMQDLEQRRPQLEELITAAQNLK NKTSNQBEARTIITDRIERI QNWDEVEQEHQNRQQLNEMLKDSTQWL EAKEEAEQVL GQARAKLESWKEGPYTVDAIQKITETKOLAKLDLRQWQTNVDVNDLALKLLRDYSAD DTRKVHMITENINASWRSIHKRVSEREALDETHRLLOQFPDLLEKFLAWLTAETTA NVLQDATRKERLLEDKSGVKELMKQWQDLQGEIEAHTDVTNLDENSQKILRSLEGSD DAVLLQRRLDNNMFKSELKSLNIRSHLEASSDQWKRLHLSQELLVWLQKDDDEL SRQAPIGGDFPAVQQNDVHRAF KRELTKKEPVMSTLETVRI FLTEQPLEGLEKLYQ EPRELPPPEERAQNVTRLRKQAEVNTWEKLNLSADWQRKIDETLERLQELQEATD ELDLKLQAEVIKGSWQPVGDLLIDSLQHLKVKALRGEIAPLKNVSHVNDLARQL TTIGIQLSPYNLSTLEDLNRWKLQVAVEDRVRLQHEAHRDFGPASQHF LSTSVQGP WERAISPKNKVPYYINHETQTTCCWDHPKMTELYQSLADLNNVRFSAYRTAMKLRRLQKA

LCLDLILSLSAACDALDQHNLIKQNDQPMIDILQIINCLTIYDRLEQEHNNLVNVPCLVD
MCLNWLNVYDTRGIRVLSFKGTIIISLKAHLEDKYRYLFKQVASSTGFCDDQRRLL
GLLLHDSIQIPROLGEVASFGGSNIPEVSRCFOFANNKPEIEAALFLDWMRLPEQSM
VLPVLJHRAAAETAKHQXCNICKCEPIIGFRYRSJXHFNYDQCOSCFPSGRVAKGH
KMHYPMVEYCTPTTSGEDVRDFAKVLKNKFRTRKYFAKHPRMGYLPVQTVLEGDNMET
PVTILNFWPVDSPASSPQLSHDDTHSRIEHYASRLAEMENSGSYLNDSPNESID
DEHLLIQHYCQSLNQDSPSPQAIILISLESEERGELERILADLEENRNQLAEY
DRLKQHEHKGLSPLSPPEMPTSPQSPRDAELIAEAKLLRQHKGRLEARMQILEDH
NKQLSOLHRLRQLLEQPOAEAKVNGTTVSSPSTSQRSDSSQPMLLRVRVGSQTSDSM
GEEDLSPPDSTSTGLEEVMBQLNNSFPSSRRGNTPKPMREDTM"
BASE COUNT 4135 a 2524 c 2876 g 2911 t
ORIGIN

Query Match 72.3%; Score 1735.6; DB 9; Length 12446;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

| | | | |
|----|------|--|------|
| QY | 628 | TGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACACTCATAGA | 687 |
| Db | 8106 | TGGAGAAGCATTCATAAAAGGGTGAGTGACGAGAGGCTGCTTTTGGAAAGAACTCATAGA | 8165 |
| QY | 688 | TTACTGCAACAGTTCCCCCTGGACCTTGGAAGATTTCTTGCCCTGGCTTACAGAAAGCTGAA | 747 |
| Db | 8166 | TTACTGCAACAGTTCCCCCTGGACCTTGGAAGATTTCTTGCCCTGGCTTACAGAAAGCTGAA | 8225 |
| QY | 748 | ACAACTGCCAATGTCTACAGGATGTACCCCGTAAGGAAAGGCTCCTAGAAGACTCCAAG | 807 |
| Db | 8226 | ACAACTGCCAATGTCTACAGGATGTACCCCGTAAGGAAAGGCTCCTAGAAGACTCCAAG | 8285 |
| QY | 808 | GGAGTAAAGAGCTGATGAAACAAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACA | 867 |
| Db | 8286 | GGAGTAAAGAGCTGATGAAACAAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACA | 8345 |
| QY | 868 | GATGTTTATACAAACCTGGATGAAACAGCCCAAAAATCCTGAGATCCCTGGAAGGTTCC | 927 |
| Db | 8346 | GATGTTTATACAAACCTGGATGAAACAGCCCAAAAATCCTGAGATCCCTGGAAGGTTCC | 8405 |
| QY | 928 | GATGATGCAGTCTCTCAACATTAGGTCOCATTTGGAAGCCAGTTCTGACCAAGTGAAGCTT | 987 |
| Db | 8406 | GATGATGCAGTCTCTCAACATTAGGTCOCATTTGGAAGCCAGTTCTGACCAAGTGAAGCTT | 8465 |
| QY | 988 | CGGAAAAAGTCTCTCAACATTAGGTCOCATTTGGAAGCCAGTTCTGACCAAGTGAAGCGT | 1047 |
| Db | 8466 | CGGAAAAAGTCTCTCAACATTAGGTCOCATTTGGAAGCCAGTTCTGACCAAGTGAAGCGT | 8525 |
| QY | 1048 | CTGCACCTTTCTCTGCAGGAACCTTCTGGTGGCTACAGCTGAAAGATGATGAATTAAGC | 1107 |
| Db | 8526 | CTGCACCTTTCTCTGCAGGAACCTTCTGGTGGCTACAGCTGAAAGATGATGAATTAAGC | 8585 |
| QY | 1108 | CGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTTCAGAACGAGAACGATGTACATAGG | 1167 |
| Db | 8586 | CGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTTCAGAACGAGAACGATGTACATAGG | 8645 |
| QY | 1168 | GCCTTCAAGAGGGAATTGAAAACTTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTA | 1227 |
| Db | 8646 | GCCTTCAAGAGGGAATTGAAAACTTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTA | 8705 |
| QY | 1228 | CGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCGAGA | 1287 |
| Db | 8706 | CGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCGAGA | 8765 |
| QY | 1288 | GAGCTGCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG | 1347 |
| Db | 8766 | GAGCTGCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG | 8825 |
| QY | 1348 | GAGGTCAATACTAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATA | 1407 |
| Db | 8826 | GAGGTCAATACTAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATA | 8885 |
| QY | 1408 | GATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAG | 1467 |
| Db | 8886 | GATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAG | 8945 |

| | | | |
|----|------|--|------|
| QY | 1468 | CTGCGCCCAAGCTGAGGTGATCAAGGGATCCTTGGCAGCCCCGTGGGGGATCTCCTCATTTGAC | 1527 |
| Db | 8946 | CTGCGCCCAAGCTGAGGTGATCAAGGGATCCTTGGCAGCCCCGTGGGGGATCTCCTCATTTGAC | 9005 |
| QY | 1528 | TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCCTCTGAAA | 1587 |
| Db | 9006 | TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCCTCTGAAA | 9065 |
| QY | 1588 | GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCATTCTGGGCACTTCAGCTC | 1647 |
| Db | 9066 | GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCATTCTGGGCACTTCAGCTC | 9125 |
| QY | 1648 | TCACCGTATAACCTCAGCACCTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTG | 1707 |
| Db | 9126 | TCACCGTATAACCTCAGCACCTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTG | 9185 |
| QY | 1708 | GCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCT | 1767 |
| Db | 9186 | GCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCT | 9245 |
| QY | 1768 | CAGCACCTTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA | 1827 |
| Db | 9246 | CAGCACCTTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA | 9305 |
| QY | 1828 | GTGCCCTACTATATCAACCACGAGACTCAAAACAACTTGCTGSGACCATCCCAAAATGACA | 1887 |
| Db | 9306 | GTGCCCTACTATATCAACCACGAGACTCAAAACAACTTGCTGSGACCATCCCAAAATGACA | 9365 |
| QY | 1888 | GAGCTCTACCACTCTTTAGCTGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCC | 1947 |
| Db | 9366 | GAGCTCTACCACTCTTTAGCTGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCC | 9425 |
| QY | 1948 | ATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTTCAGCTGCA | 2007 |
| Db | 9426 | ATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTTCAGCTGCA | 9485 |
| QY | 2008 | TGTGATGCCCTTGGACCAGCAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAG | 2067 |
| Db | 9486 | TGTGATGCCCTTGGACCAGCAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAG | 9545 |
| QY | 2068 | ATTATTAAATGTTTGACCACCTATTATGACCGCTGGAGCAAGAGCACAAATTTGGTC | 2127 |
| Db | 9546 | ATTATTAAATGTTTGACCACCTATTATGACCGCTGGAGCAAGAGCACAAATTTGGTC | 9605 |
| QY | 2128 | AACGTCCCTCTCTGCGTGGATATGTGTCTGAACCTGGCTGGTGAATGTTATGATACGGGA | 2187 |
| Db | 9606 | AACGTCCCTCTCTGCGTGGATATGTGTCTGAACCTGGCTGGTGAATGTTATGATACGGGA | 9665 |
| QY | 2188 | CGAACAGGAGGATCCGTGTCTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA | 2247 |
| Db | 9666 | CGAACAGGAGGATCCGTGTCTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA | 9725 |
| QY | 2248 | CATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGT | 2307 |
| Db | 9726 | CATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGT | 9785 |
| QY | 2308 | GACCAGCGCAGGCTGGGCTCTCTCTGTCATGATTTCTATCCAAATTCCAAGACAGTTGGGT | 2367 |
| Db | 9786 | GACCAGCGCAGGCTGGGCTCTCTCTGTCATGATTTCTATCCAAATTCCAAGACAGTTGGGT | 9845 |
| QY | 2368 | GAAGTTGCATCCTTTGGGGGCGAGTAACATTGAGC | 2401 |
| Db | 9846 | GAAGTTGCATCCTTTGGGGGCGAGTAACATTGAGC | 9879 |

| | | | |
|------------|--------------------------------------|----------|-----------------|
| RESULT 8 | | | |
| AX409637 | Sequence 2284 from Patent WO0229103. | | |
| LOCUS | AX409637 | 13957 bp | DNA |
| DEFINITION | AX409637 | linear | PAT 14-JUN-2002 |
| ACCESSION | AX409637.1 | | |
| VERSION | GI:21442342 | | |
| KEYWORDS | | | |
| SOURCE | Homo sapiens (human) | | |

ACCESSION AX538581
VERSION AX538581.1 GI:25271086
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 1 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
source Location/Qualifiers
1.13957
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 4602 a 2781 c 3122 g 3452 t
ORIGIN
Query Match 72.3%; Score 1735.6; DB 6; Length 13957;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 628 TGGGCAACATCTGTAGATGGACAGAAGACCGCTGGTTCCTTTTACAAGACACTCATAGA 687
DB 8216 TGGAGAGCATTCTATAAAGGGTGTGCTGAGCGAGAGGCTGCTTTGGAAAGAACTCATAGA 8275
QY 688 TTACTGCAACAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGGCTGGCTTACAGAAGCTGAA 747
DB 8276 TTACTGCAACAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGGCTGGCTTACAGAAGCTGAA 8335
QY 748 ACAACTGCCAATGTCTACAGGATGCTACCCGTAAAGGAAGCTCCTAGAAGACTCCAAG 807
DB 8336 ACAACTGCCAATGTCTACAGGATGCTACCCGTAAAGGAAGCTCCTAGAAGACTCCAAG 8395
QY 808 GGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAAGTGAAATTGAAGCTCACACA 867
DB 8396 GGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAAGTGAAATTGAAGCTCACACA 8455
QY 868 GATGTTTATCACAACTGGATGAAACAGCCAAAAAATCCTGAGATCCCTGGAGGTTCC 927
DB 8456 GATGTTTATCACAACTGGATGAAACAGCCAAAAAATCCTGAGATCCCTGGAGGTTCC 8515
QY 928 GATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTT 987
DB 8516 GATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTT 8575
QY 988 CGGAAAAAGTCTCTCAACATTAGTCCCATTTTGGAGCCAGTCTGACCAAGTGGAGGCGT 1047
DB 8576 CGGAAAAAGTCTCTCAACATTAGTCCCATTTTGGAGCCAGTCTGACCAAGTGGAGGCGT 8635
QY 1048 CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 1107
DB 8636 CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 8695
QY 1108 CGGCAGGCACCTATTGGAGGCGGACTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGG 1167
DB 8696 CGGCAGGCACCTATTGGAGGCGGACTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGG 8755
QY 1168 GCCTTCAAGAGGGGAATTGAAAACCTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTA 1227
DB 8756 GCCTTCAAGAGGGGAATTGAAAACCTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTA 8815
QY 1228 CGAATATTCTGACAGAGCAGCGCTTTTGGAGGACTAGAGAACTCTACAGGAGCCCCAGA 1287
DB 8816 CGAATATTCTGACAGAGCAGCGCTTTTGGAGGACTAGAGAACTCTACAGGAGCCCCAGA 8875
QY 1288 GAGCTGCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 1347
DB 8876 GAGCTGCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 8935
QY 1348 GAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 1407

DB 8936 GAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 8995
QY 1408 GATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCCACCGATGAGCTGGACCTCAAG 1467
DB 8996 GATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCCACCGATGAGCTGGACCTCAAG 9055
QY 1468 CTGCGCCCAAGCTGAGGTGATCAAGGGATCCTTGGCAGCCCCGTGGCGGATCTCCTCATTTGAC 1527
DB 9056 CTGCGCCCAAGCTGAGGTGATCAAGGGATCCTTGGCAGCCCCGTGGCGGATCTCCTCATTTGAC 9115
QY 1528 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAGAAATTGCGCCTCTGAAA 1587
DB 9116 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAGAAATTGCGCCTCTGAAA 9175
QY 1588 GAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTTCAGCTC 1647
DB 9176 GAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTTCAGCTC 9235
QY 1648 TCACCGTATAAACCCTCAGCACCTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTG 1707
DB 9236 TCACCGTATAAACCCTCAGCACCTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTG 9295
QY 1708 GCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGAGACTTTTGGTCCAGCATCT 1767
DB 9296 GCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGAGACTTTTGGTCCAGCATCT 9355
QY 1768 CAGCACCTTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAA 1827
DB 9356 CAGCACCTTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAA 9415
QY 1828 GTGCCCTACTATATCAACCACGAGACTCAAAACAACCTTGTGGGACCACATCCCAAAATGACA 1887
DB 9416 GTGCCCTACTATATCAACCACGAGACTCAAAACAACCTTGTGGGACCACATCCCAAAATGACA 9475
QY 1888 GAGCTCTACCAAGTCTTTAGCTGACCTGAATAATGTTCAGATTTCTCAGCTTATAGGACTGCC 1947
DB 9476 GAGCTCTACCAAGTCTTTAGCTGACCTGAATAATGTTCAGATTTCTCAGCTTATAGGACTGCC 9535
QY 1948 ATGAACCTCCGAAGACTGCAGAGGCCCCCTTTGCTTGGATCTCTTGAGCCTGTTCAGCTGCA 2007
DB 9536 ATGAACCTCCGAAGACTGCAGAGGCCCCCTTTGCTTGGATCTCTTGAGCCTGTTCAGCTGCA 9595
QY 2008 TGTGATGCCCTTGGACCAACACCTCAAGCAAAATGACAGCCCATGGATATCCTTGCAG 2067
DB 9596 TGTGATGCCCTTGGACCAACACCTCAAGCAAAATGACAGCCCATGGATATCCTTGCAG 9655
QY 2068 ATTATTAATTGTTTGACCACTATTTATGACCGCTTGAACCTGGGCTGCTGAATGTTTATGATACGGGA 2127
DB 9656 ATTATTAATTGTTTGACCACTATTTATGACCGCTTGAACCTGGGCTGCTGAATGTTTATGATACGGGA 9715
QY 2128 AACGTCCTCTCTGCGTGGATATGTCTGAACCTGGGCTGCTGAATGTTTATGATACGGGA 2187
DB 9716 AACGTCCTCTCTGCGTGGATATGTCTGAACCTGGGCTGCTGAATGTTTATGATACGGGA 9775
QY 2188 CGAACAGGGAGGATCCGTCCTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 2247
DB 9776 CGAACAGGGAGGATCCGTCCTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 9835
QY 2248 CATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTGT 2307
DB 9836 CATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTGT 9895
QY 2308 GACCAGCGCAGGCTGGGCTCTCTGATGATTTCTATCCAAATTCGAAGACAGTTGGGT 2367
DB 9896 GACCAGCGCAGGCTGGGCTCTCTGATGATTTCTATCCAAATTCGAAGACAGTTGGGT 9955
QY 2368 GAAGTTGCATCCTTTGGGGGCGAGTAACATTGAGC 2401
DB 9956 GAAGTTGCATCCTTTGGGGGCGAGTAACATTGAGC 9989

RESULT 10

Db 8456 GATGTTTATCACAACTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGTTCC 8515

Qy 928 GATGATGCAGTCCTGTTACAAAGACGTTTGGATAAACATGAACCTTCAAGTGGAGTGAACCT 987

Db 8516 GATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCT 8575

Qy 988 CGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCGT 1047

Db 8576 CGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCGT 8635

Qy 1048 CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 1107

Db 8636 CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 8695

Qy 1108 CGGCAGGCACCTTATGGAGGCGACTTTCACAGCAGTTTCAGAAAGCAGACGATGTACATAGG 1167

Db 8696 CGGCAGGCACCTTATGGAGGCGACTTTCACAGCAGTTTCAGAAAGCAGACGATGTACATAGG 8755

Qy 1168 GCCTTCAAGAGGGAAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTA 1227

Db 8756 GCCTTCAAGAGGGAAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTA 8815

Qy 1228 CGAATATTTCTGACAGAGCAGCCTTTTGGAGGACTAGAGAAACTCTACAGGAGCCCAAG 1287

Db 8816 CGAATATTTCTGACAGAGCAGCCTTTTGGAGGACTAGAGAAACTCTACAGGAGCCCAAG 8875

Qy 1288 GAGCTGCCCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 1347

Db 8876 GAGCTGCCCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 8935

Qy 1348 GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 1407

Db 8936 GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 8995

Qy 1408 GATGAGACCCCTTGAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAG 1467

Db 8996 GATGAGACCCCTTGAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAG 9055

Qy 1468 CTGCGCCCAAGTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGAC 1527

Db 9056 CTGCGCCCAAGTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGAC 9115

Qy 1528 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGGCCCTCTGAAA 1587

Db 9116 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGGCCCTCTGAAA 9175

Qy 1588 GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACACCTTTGGGCATTCAGCTC 1647

Db 9176 GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTC 9235

Qy 1648 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTG 1707

Db 9236 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTG 9295

Qy 1708 GCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGGACTTTGGTCCAGCATCT 1767

Db 9296 GCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGGACTTTGGTCCAGCATCT 9355

Qy 1768 CAGCACTTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 1827

Db 9356 CAGCACTTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 9415

Qy 1828 GTGCCCTACTATATCAACCACGAGACTCAAAACAACTTGCTGGAGACCATCCCAAAATGACA 1887

Db 9416 GTGCCCTACTATATCAACCACGAGACTCAAAACAACTTGCTGGAGACCATCCCAAAATGACA 9475

Qy 1888 GAGCTCTACCACTCTTAGCTGACCTGAATAATGTGAGATTTCTCAGCTTATAGGACTGCC 1947

Db 9476 GAGCTCTACCACTCTTAGCTGACCTGAATAATGTGAGATTTCTCAGCTTATAGGACTGCC 9535

Qy 1948 ATGAAACTCCGAAGACTGCAGAAGGCCCTTTGGCTTGGATCTCTTGAGCCTGTGAGCTGCA 2007

Db 9536 ATGAAACTCCGAAGACTGCAGAAGGCCCTTTGGCTTGGATCTCTTGAGCCTGTGAGCTGCA 9595

Qy 2008 TGTGATGCCCTTGGACCAGCACAACTCAAGCAAAAATGACCAGCCCATGGATATCCTGCAG 2067

Db 9596 TGTGATGCCCTTGGACCAGCACAACTCAAGCAAAAATGACCAGCCCATGGATATCCTGCAG 9655

Qy 2068 ATTATTAATTGTTGACCACACTATTATGACCGCCTGGAGCAAGAGACAAACAATTTGGTC 2127

Db 9656 ATTATTAATTGTTGACCACACTATTATGACCGCCTGGAGCAAGAGACAAACAATTTGGTC 9715

Qy 2128 AACGTCCCTCTCTGCGTGGATATGTCTGAACCTGGCTGGTGAATGTTTATGATACGGGA 2187

Db 9716 AACGTCCCTCTCTGCGTGGATATGTCTGAACCTGGCTGGTGAATGTTTATGATACGGGA 9775

Qy 2188 CGAACAGGGAGGATCCGTGTCTCTGCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCA 2247

Db 9776 CGAACAGGGAGGATCCGTGTCTCTGCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCA 9835

Qy 2248 CATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGT 2307

Db 9836 CATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGT 9895

Qy 2308 GACCAGCGCAGCTGGCCCTCCTTCTGTCATGATTTCTATCCAAATTTCCAAAGACAGTTGGGT 2367

Db 9896 GACCAGCGCAGCTGGCCCTCCTTCTGTCATGATTTCTATCCAAATTTCCAAAGACAGTTGGGT 9955

Qy 2368 GAAGTTGCATCCTTTGGGGGCAGTAACATTGAGC 2401

Db 9956 GAAGTTGCATCCTTTGGGGGCAGTAACATTGAGC 9989

RESULT 11

AX538620

LOCUS AX538620 Sequence 40 from Patent WO0229056. linear PAT 23-NOV-2002

DEFINITION AX538620

ACCESSION AX538620

VERSION AX538620.1 GI:25271166

KEYWORDS synthetic construct

SOURCE synthetic construct

ORGANISM synthetic construct

REFERENCE 1 artificial sequences.

AUTHORS Chamberlain, J.S. and Harper, S.Q.

TITLE Mini-dystrophin nucleic acid and peptide sequences

JOURNAL Patent: WO 0229056-A 40 11-APR-2002;

THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

FEATURES

Location/Qualifiers

source

1. .5339

/organism="synthetic construct"

/mol_type="genomic DNA"

/db_xref="taxon:32630"

/note="Synthetic"

BASE COUNT 1638 a 1191 c 1187 g 1323 t

ORIGIN

Query Match 71.8%; Score 1725; DB 6; Length 5339;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1721; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 681 TCATAGATTACTGCAACAGTTCCCTCGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGA 740

Db 1546 TCATAGATTACTGCAACAGTTCCCTCGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGA 1605

Qy 741 AGCTGAAACAACTGCCAATGTCTTACAGGATGTCTACCCGTAAGGAAAGGCTCCTTAGAGA 800

Db 1606 AGCTGAAACAACTGCCAATGTCTTACAGGATGTCTACCCGTAAGGAAAGGCTCCTTAGAGA 1665

Qy 801 CTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 860

Db 1666 CTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 1725

Qy 861 TCACACAGATGTTTATCAACAACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGA 920

Db 1726 TCACACAGATGTTTATCAACAACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGA 1785

QY 921 AGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTTGGATAACATGAACCTTCAAGTGGAG 980
Db 1786 AGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTTGGATAACATGAACCTTCAAGTGGAG 1845
QY 981 TGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAAGTG 1040
Db 1846 TGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAAGTG 1905
QY 1041 GAAGCGCTGACACCTTTCTCTGACGAACTTCTGGTGCGGTACAGCTGAAAGATGATGA 1100
Db 1906 GAAGCGCTGACACCTTTCTCTGACGAACTTCTGGTGCGGTACAGCTGAAAGATGATGA 1965
QY 1101 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCGATTTGGAAGCGAAGCAACGATGT 1160
Db 1966 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCGATTTGGAAGCGAAGCAACGATGT 2025
QY 1161 ACATAGGGCTTCAAGAGGGAATTGAAACTAAAGAACCTGTGAATCATGAGTACTCTTGA 1220
Db 2026 ACATAGGGCTTCAAGAGGGAATTGAAACTAAAGAACCTGTGAATCATGAGTACTCTTGA 2085
QY 1221 GACTGTACGAATATTTCTGACAGAGCGCCCTTTGGAAGGACTAGAGAAACTCTACCAAGGA 1280
Db 2086 GACTGTACGAATATTTCTGACAGAGCGCCCTTTGGAAGGACTAGAGAAACTCTACCAAGGA 2145
QY 1281 GCCCAGAGAGTGCCTCCTGAGGAGAGAGCCCGAGAAATGTCATCGGCTTCTACGAAAGCA 1340
Db 2146 GCCCAGAGAGTGCCTCCTGAGGAGAGAGCCCGAGAAATGTCATCGGCTTCTACGAAAGCA 2205
QY 1341 GGCTGAGGAGTCAATACTAGTGGGAAAAAATTGAACCTGCACTCGCTGACTGGCAGAG 1400
Db 2206 GGCTGAGGAGTCAATACTAGTGGGAAAAAATTGAACCTGCACTCGCTGACTGGCAGAG 2265
QY 1401 AAAAATAGATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 1460
Db 2266 AAAAATAGATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 2325
QY 1461 CCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCT 1520
Db 2326 CCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCT 2385
QY 1521 CATTGACTCTCTCAAGATCACCTCGAGAAAGTCAAGGCACCTCGAGGAGAAATTGGGCC 1580
Db 2386 CATTGACTCTCTCAAGATCACCTCGAGAAAGTCAAGGCACCTCGAGGAGAAATTGGGCC 2445
QY 1581 TCTGAAAGAGAACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGSCAT 1640
Db 2446 TCTGAAAGAGAACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGSCAT 2505
QY 1641 TCAGCTCTCACCGTATAACCTCAGCACCTTGGAAGACCTGAACACACAGATGGAAGCTTCT 1700
Db 2506 TCAGCTCTCACCGTATAACCTCAGCACCTTGGAAGACCTGAACACACAGATGGAAGCTTCT 2565
QY 1701 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGACCTTTGGTCC 1760
Db 2566 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGACCTTTGGTCC 2625
QY 1761 AGCATCTCAGCACCTTCTTTCCACGTCGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCC 1820
Db 2626 AGCATCTCAGCACCTTCTTTCCACGTCGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCC 2685
QY 1821 AAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAACAACTTGCTGGGACCATCCCAA 1880
Db 2686 AAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAACAACTTGCTGGGACCATCCCAA 2745
QY 1881 AATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAATGTCAGATTTCTCAGCTTATAG 1940
Db 2746 AATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAATGTCAGATTTCTCAGCTTATAG 2805
QY 1941 GACTGCCATGAAACTCCGAAGACTGCAGAGGCCCTTTTGCTTGGATCTCTTGAGCCCTGTC 2000
Db 2806 GACTGCCATGAAACTCCGAAGACTGCAGAGGCCCTTTTGCTTGGATCTCTTGAGCCCTGTC 2865

QY 2001 AGCTGCATGTGATGCCTTGGACCAGACAAACCTCAAGCAAAATGACCAGCCCATGGATAT 2060
Db 2866 AGCTGCATGTGATGCCTTGGACCAGACAAACCTCAAGCAAAATGACCAGCCCATGGATAT 2925
QY 2061 CCTGCAGATTATTAATTGTTTGACCACTATTTATGACCCGCTGGAGCAAGACAAACAA 2120
Db 2926 CCTGCAGATTATTAATTGTTTGACCACTATTTATGACCCGCTGGAGCAAGACAAACAA 2985
QY 2121 TTTGTTCAACGTCCTCTCTCGGTGGATATGHTCTGAACTGGCTGCTGAATGTTTATGA 2180
Db 2986 TTTGTTCAACGTCCTCTCTCGGTGGATATGHTCTGAACTGGCTGCTGAATGTTTATGA 3045
QY 2181 TACGGGACGAACAGGAGGATCCGTCTCTGCTCTTTTAAAACTGGCATCATTTCCCTGTG 2240
Db 3046 TACGGGACGAACAGGAGGATCCGTCTCTGCTCTTTTAAAACTGGCATCATTTCCCTGTG 3105
QY 2241 TAAAGCACATTTGGAAGACAAAGTACAGATACCTTTTCAAGCAAAGTGGCAAGTTCAACAGG 2300
Db 3106 TAAAGCACATTTGGAAGACAAAGTACAGATACCTTTTCAAGCAAAGTGGCAAGTTCAACAGG 3165
QY 2301 ATTTTGTGACCAGCGCAGGCTGGGCCTCCTTCTGATGATTTATCCAAATTCCAAGACA 2360
Db 3166 ATTTTGTGACCAGCGCAGGCTGGGCCTCCTTCTGATGATTTATCCAAATTCCAAGACA 3225
QY 2361 GTTGGTGAAAGTTGCATCCTTTGGGGCAGTAACATTGAGC 2401
Db 3226 GTTGGTGAAAGTTGCATCCTTTGGGGCAGTAACATTGAGC 3266

RESULT 12
AR220819 13977 bp DNA linear PAT 26-SEP-2002
LOCUS Sequence 60 from patent US 6426186.
DEFINITION AR220819
ACCESSION AR220819
VERSION AR220819.1 GI:23327696
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 13977)
AUTHORS Jones,K.A., Volkmut,W. and Walker,M.G.
TITLE Bone remodeling genes
JOURNAL Patent: US 6426186-A 60 30-JUL-2002;
FEATURES Location/Qualifiers
source 1..13977
BASE COUNT 4596 a 2765 c 3120 g 3453 t 43 others
ORIGIN

Query Match 71.8%; Score 1724.6; DB 6; Length 13977;
Best Local Similarity 98.6%; Pred.No. 0;
Matches 1750; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

QY 628 TGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACACTCATAGA 687
Db 8216 TGGAGAAGCAATTCATAAAGGGTGACTGAGCGAGAGGCTGCTTTGGAAGAAACTCATAGA 8275
QY 688 TTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGGCTGGCTTACAGAAGCTGAA 747
Db 8276 TTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGGCTGGCTTACAGAAGCTGAA 8335
QY 748 ACAACTGCCAATGTCTTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAAG 807
Db 8336 ACAACTGCCAATGTCTTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAAG 8395
QY 808 GGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACA 867
Db 8396 GGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACA 8455
QY 868 GATGTTTATCAACACCTGGATGAAAACAGCCAAAAAATCTTGAGATCCCTGGAAGGTTCC 927
Db 8456 GATGTTTATCAACACCTGGATGAAAACAGCCAAAAAATCTTGAGATCCCTGGAAGGTTCC 8515

QY 928 GATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAGTGAACCTT 987
Db 8516 GATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAGTGAACCTT 8575
QY 988 CGGAAAAAGTCTCTCAACATTAGGTCCTCCATTTGGAAGCCAGTTCTGACCAAGTGGAAAGCGT 1047
Db 8576 CGGAAAAAGTCTCTCAACATTAGGTCCTCCATTTGGAAGCCAGTTCTGACCAAGTGGAAAGCGT 8635
QY 1048 CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 1107
Db 8636 CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 8695
QY 1108 CGGCAGGACCTTATTGGAGGCGACTTTCCAGCAGTTTCAAGAGCAAGCATGTACATAGG 1167
Db 8696 CGGCAGGACCTTATTGGAGGCGACTTTCCAGCAGTTTCAAGAGCAAGCATGTACATAGG 8755
QY 1168 GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTA 1227
Db 8756 GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTA 8815
QY 1228 CGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACTCTACCAAGAGCCCAAGA 1287
Db 8816 CGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACTCTACCAAGAGCCCAAGA 8875
QY 1288 GAGCTGCCCTCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 1347
Db 8876 GAGCTGCCCTCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 8935
QY 1348 GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 1407
Db 8936 GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 8995
QY 1408 GATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAG 1467
Db 8996 GATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAG 9055
QY 1468 CTGCGCCCAAGCTGAGGTGATCAAGGGATCCTTGGCAGCCCGTGGCGGATCTCCTCATTCAC 1527
Db 9056 CTGCGCCCAAGCTGAGGTGATCAAGGGATCCTTGGCAGCCCGTGGCGGATCTCCTCATTCAC 9115
QY 1528 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAAAATTGGCCTCTGAAA 1587
Db 9116 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAAAATTGGCCTCTGAAA 9175
QY 1588 GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACACATTTGGGCATTCAGCTC 1647
Db 9176 GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACACATTTGGGCATTCAGCTC 9235
QY 1648 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAAGTG 1707
Db 9236 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAAGTG 9295
QY 1708 GCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGCACTTTGGTCCAGCATCT 1767
Db 9296 GCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGCACTTTGGTCCAGCATCT 9355
QY 1768 CAGCACTTTCTTCCACGTCTGTCCAGGTCCTCGGAGAGAGCCATCTCGCCAAACAAA 1827
Db 9356 CAGCACTTTCTTCCACGTCTGTCCAGGTCCTCGGAGAGAGCCATCTCGCCAAACAAA 9415
QY 1828 GTGCCCTACTATATCAACACCGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACA 1887
Db 9416 GTGCCCTACTATATCAACACCGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACA 9475
QY 1888 GAGCTCTACCACTTTAGCTGACCTGTAATATGTGATGATTTCTCAGCTTATAGGACTGCC 1947
Db 9476 GAGCTCTACCACTTTAGCTGACCTGTAATATGTGATGATTTCTCAGCTTATAGGACTGCC 9535
QY 1948 ATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGCA 2007
Db 9536 ATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGCA 9595
QY 2008 TGTGATGCCCTTGGACCAACCAACCTCAAGCAAAAATGACCAGCCCATGGATATCCTGCGAG 2067

Db 9596 TGTGATGCCCTTGGACCAACCAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCGAG 9655
QY 2068 ATTATTAAATTGTTGACCACACTATTATGACCGCCTGGAGCAAGAGACAAACAATTTGGTC 2127
Db 9656 ATTATTAAATTGTTGACCACACTATTATGACCGCCTGGAGCAAGAGACAAACAATTTGGTC 9715
QY 2128 AACGTCCCTCTCTGCGTGGATATGTGCTGAACTGGCTGGCTGCTGAATGTTTATGATACGGGA 2187
Db 9716 AACGTCCCTCTCTGCGTGGATATGTGCTGAACTGGCTGGCTGCTGAATGTTTATGATACGGGA 9775
QY 2188 CGAACAGGGAGGATCCGTGTCTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCA 2247
Db 9776 CGAACAGGGAGGATCCGTGTCTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCA 9835
QY 2248 CATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACACAGGATTTTGT 2307
Db 9836 CATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACACAGGATTTTGT 9895
QY 2308 GACCAGCGCAGCT-GGGCCTCCTTCTGTCATGATTTCTATCCAATTCACAGACAGTTGGG 2366
Db 9896 GACCAGCGCAGCTGGGGCCTCCTTCTGTCATGATTTCTATCCAATTCACAGACAGTTGGG 9955
QY 2367 TGAAGTTGCATCCTTTGGGGGCAGTAACATTGAGC 2401
Db 9956 TGAAGTTGCATCCTTTGGGGGCAGTAACATTGAGC 9990

RESULT 13
AF070485
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
5'UTR
CDS

AF070485 13887 bp mRNA linear MAM 09-DEC-1998
Canis familiaris dystrophin mRNA, complete cds.
AF070485
AF070485.1 GI:3982750
Canis familiaris (dog)
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 13887)
Carville,K.S., Mann,C.J., Schatzberg,S.J. and Wilton,S.D.
Direct Submission
Submitted (04-JUN-1998) ANRI, Pathology, University of Western
Australia, Verdun Street, Nedlands, WA 6018, Australia
Location/Qualifiers
1. 13887
/organism="Canis familiaris"
/mol_type="mRNA"
/strain="breed Golden retriever"
/db_xref="taxon:9615"
<1. .289
290. .11332
/codon_start=1
/product="dystrophin"
/protein_id="AAC83646.1"
/db_xref="GI:3982751"
/translation="MLWWEVEEDCYEREDVQKTFKWNVAQFSKFGKQHIENLFSDL
QDGRLLDLLEGLTGQKLPKEKGS TRVHALNNVNKALRVLQKNVDLVNIGSTDIVDG
NHKLTGLIWNILHWQVKNVMKNIMAGLQQTNSEKILLSWVRQSTRNYPQVNVINFT
TSWSDGLALNALIHSHRPDLFDWNSVVCQSQSATORLEHAFNIAKYQLGIEKLDPEV
ATTYPDKKSILMYITSLFQVLPQVSI EAIQIEVEMLP RPSQVTRREEHQIHHQMHVSQ
QITVSLAQGYERAPSPFRPKFSYAYTAAAYVTTSDPTRSP LPSQHLTEPEDKSFGRS
LTETEANLDSYQTAL EEEVL SWLSAEDALQAQGEI SNDVEEVKEQFHTEHGYNMDLTS
HGRVGNVLQGLSQLTGKLS EDETEVQEQMNLNLSRWECLRVASMEKQSNLHKVL
MDLQNLQKELNDWLTKEERTKMEKEPLGPDIEDLKRQVQKHVLQEDLEQEQRV
NSLTHMVVVVDES GDHATAALEEQKVLGDRANICRWTEDRWVLLQDILLKWRFT
EEQLFSAWLSEKEDAVNKIHTTGFKDQSEVLNQLKLA VLKTDLEKKKQTMDKLCSL
METVMVTTREHILVKHAQEELPPPPQKKRQIIVDSEIRKRLDVIDITELHSWITRSE
AVLQSPFAIYRKEGNSDLKEKNVIEREKEKFRKLQDASRSQAALVEQMVNNEGYN
ADSIKQASEQLNSRWTFECQLLSERLNWLEYQNNIITFYNQLQQLEQMTTAEWNLKT
OPTTSEPTAIKSQLKICKDEINRLSALOPQIERLKIQSI ALKEKGQGMFLDADFVA
FTNHFNOVFADVQAREKELQTI FDSLPPMRYQETMSTILTWTWQQSETKLSIPQVTVTE

| | | | | | | | | | |
|---|--------|---|---------|------------|------|--------|--------|------|----|
| YDIMEQRLGELQALQSSLOEQOQNGNLNLSSTTVKEMSKKAPLSDISRKYQSEFEIEGR | | | | | | | | | |
| WKLSSQLVEHCQKLEEQMAKLRKIQNHIKTLKWKITEVDVFLKEEWPALGDSIILKR | | | | | | | | | |
| OLKQCRLLVNDIQIOPSLSNVNEGAKQMKNEAEPEFAGRLTELRELTQWDMCRO | | | | | | | | | |
| VYARKEALKGLDKTVSLQKDSLSEMEHWMTQAEIEYKLELDFTYKTPDELQTAENGMR | | | | | | | | | |
| AKEEAQCKEAKVKLLTESVNSVIAQAPPAQAEALKKELDTLTNNQWLCTRLNGCKT | | | | | | | | | |
| LEEYWAACWHELLSYLEKANKWLSEVEVKLTITTENISGGAEEIAEVLDSLENLMQHS | | | | | | | | | |
| NPNOIRILAQTLTDGVMDELINEELETENSRRWELHEEAVRRQKLEQISQSAQIE | | | | | | | | | |
| KSLHLIOESLSSIDKQLAAYIADKVDAQAQMPQEAQIQSDLSHSHEISLEEMKKHQGK | | | | | | | | | |
| ETAQRVLSQIDVAQKKLQDVSMKFRFLQKPAQNEQRLQESKMILDEVKMLPALETKS | | | | | | | | | |
| VEQEYVQSQNLHCNVNLYKLSSEVKSEVEMVIKTGRIVQKKQTNPENKELDERVTALKL | | | | | | | | | |
| HYNELGAKVTERKQOLEKCLKLSRKMVKEMNALTEWLAATDMELTKRSAYEGMPSNLD | | | | | | | | | |
| SEVANGKATQKEIEKQKVHLKSVTEVGEALKTVLGKXKMLVEDKLSLNSNWIATVSR | | | | | | | | | |
| AEEWNLLELEYQKHMETFDQNVYITNWIIQADALLDSEKKPQOKKEDILKRLKAEW | | | | | | | | | |
| NDIRPKVDSRDOQAANLMANRGDHRKRVPEPKISELNRHFAAISHRIKTKGASIPLEK | | | | | | | | | |
| LEQFNSDIQKLLPLEAEIOQVNLKEEDFNKMSSEDNKGVKELLQKQNLQWIK | | | | | | | | | |
| ERKEBEIKIKQOLLQTKHNLKDLRSQRRKKALEISHQWYQKRAADDLLKLDIEK | | | | | | | | | |
| KLASLEPRDERKIKEIDRELQKKKEELNAVRRQAEGLSEDGAAMAVEPTQIQLSKRW | | | | | | | | | |
| REIESFPAQFRRINFQAIHTVHEESVWAMTEDMPLEISYVPSYLTITEITHVSQALSEV | | | | | | | | | |
| EELNAPDLCAQDFEDLFKQESLNKIDSLQIQISGRIDIHNKKTAAHLSATPAERA | | | | | | | | | |
| KLQEAUSRLDFOWERVNNMYKDRQGRFDRSVEKWRFRPHYDMKILNQWLTEAEQFLKKT | | | | | | | | | |
| QIPENWEHAKYKYLKELQDGIQGRQSVVRVNLATGEEIIQOSSKTDAASILQEKLSL | | | | | | | | | |
| NLRWQEVCKQLAERKRLLEEQKNILSEFQRDVNEFVLWEADNVANIPLRPGNEQQL | | | | | | | | | |
| KEXLEQVKLLAEELPLRQGLKQLNETGGTVLSAPLSPEEQDKLENKQLQTNLQWIK | | | | | | | | | |
| VSRNLPKQEIEIEAHVKDLQLEELNHLHLLWLSPRNQLEIYNQPNQTPGPDIEIE | | | | | | | | | |
| VAVQAKQPDVEGILSKGQHLKYEKPATQPAKRKLELSDSSDWKVVITQLLELRAKQPGP | | | | | | | | | |
| APGLTTVRAPSQTVTLVTPATVKETAISKLEMPSSILLLEVPALEFNRAWTELTDW | | | | | | | | | |
| LSLLDRVYKSRQVMVGDDLEINEMIIKQKATLQDLQRRPQLEELITAAQNLKNKTSN | | | | | | | | | |
| QEARTIITDRIERIQSQWDEVEQHLQNRRLQLTEMLKDSQWLEAKEEAEQVLQARA | | | | | | | | | |
| KLESWEAPYTVDAIQKKITETKQAKDLRQWQINVDVANDLALKLTDYSADDTRKV | | | | | | | | | |
| HMITENINASWASIHKRLSEREALEETHRLLQOFPDLLEKFLAWLTEAETTANVLQD | | | | | | | | | |
| ATHKERLLEDSKGVRELKMQWDLQGEIEAHTDIYHNLDENGQKVLRSLEGSDDAALL | | | | | | | | | |
| QRRLDNNFKNSELRKKSLSNIRSHLEASSDQWKRHLHLSLQELLVWLQKDDLSRQAP | | | | | | | | | |
| IGDDFFPAQKQNDVHRAFKRELKTEKPEVIMSTLETVRIFLTEQPLEGLEKLYQEPREL | | | | | | | | | |
| PPEERAQNVTRLLRKQAEVNTQWEKLNHVSADWQKIDEALERLOELQEADELDLKL | | | | | | | | | |
| LRQAEVTKGSQWPVGDLILDSLDQHLKVKALRGEITPLKENYSYNDLARQLTTLGI | | | | | | | | | |
| QLSPYNLTLEDLNRWKLQVAIEDRIQLHEARDFGPASQHFSLTSVQGPWERAI | | | | | | | | | |
| SPNKVPYINHETQTTCDWHPKMTELYQSLADLNNVRFPSAYRTAMKLRRLQKALCLDL | | | | | | | | | |
| LSLSAADALDOHNLKQNDQPMILQVINCLTTIYDRLEQEHNNLVNPLCVDMLNW | | | | | | | | | |
| LLNVYDTGRGTRIRVLSFKGTIISLKAHLEDKRYLFLKQVASSSTGFCDDORLGLLH | | | | | | | | | |
| DSIQIPQLGEVASFGGSNIEPSVRSQFPANNKPEIEAALFLDMRLEPQSMVWLVP | | | | | | | | | |
| LHRVAAEETAKHOAKNCNICECPIIGFRYSRLKHFNDIQCSCFFSGRVAKGHKHYP | | | | | | | | | |
| MVEYCTPTSGEDVRDFAKVLKNKFKTRKYFAKHPRMGVLPVQTVLEGDNMETPVTLL | | | | | | | | | |
| NFWPVDSPASSPOLSHDDTHSRIEHYASRLKMNENSGSYLNDISIPNESIDDEHLL | | | | | | | | | |
| IQHYWRSLNQESPLSPRSPAQILISLESEERGELEIRILADLEGRRNRLQAEYDRLKQ | | | | | | | | | |
| QHEHKGLSPLSPPEMMPSTSPQSPRDAELIAEAKLRQKHQLEARMQILEDHINKOLE | | | | | | | | | |
| SQLHRLQLLEQPOAEAKVNGTIVSSPSTLSQRSDSQPMLLRVVGSQTSSEMGEDL | | | | | | | | | |
| LSPPDQSTSTGLEEVMEQLNHSFPSSRGRNTPGKPMREDTM" | | | | | | | | | |
| 11333. >13887 | | | | | | | | | |
| 11789..11828 | | | | | | | | | |
| /rpt_type=tandem | | | | | | | | | |
| /rpt_unit=ca | | | | | | | | | |
| BASE COUNT | 4455 a | 2906 c | 3164 g | 3362 t | | | | | |
| ORIGIN | | | | | | | | | |
| Query Match | 64.2%; | Score | 1540.4; | DB | 4; | Length | 13887; | | |
| Best Local Similarity | 91.8%; | Pred. No. | 0; | | | | | | |
| Matches | 1628; | Conservative | 0; | Mismatches | 146; | Indels | 0; | Gaps | 0; |
| Qy | 628 | TGGGCAAAACATCTGTAGATGGACAGAAAGACCGCTGGGTTCTTTTACAAAGACACTCATAGA | 687 | | | | | | |
| Db | 8282 | TGGGCAAGCATCCATAAAAGATTGAGTGAGCGAGAGGCTGCTCTGGAAGAAACCCACAGA | 8341 | | | | | | |
| Qy | 688 | TTACTGCAACAGTTC | 747 | | | | | | |
| Db | 8342 | TTACTGCAACAGTTC | 8401 | | | | | | |
| Qy | 748 | ACAACTGCCAATGTCTCAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAAGACTCCAAAG | 807 | | | | | | |
| Db | 8402 | ACAACTGCCAACGCTCTGCAGGATGCCACCCATAAGGAAAGGCTTCTAGAAAGATTCCAAAG | 8461 | | | | | | |
| Qy | 808 | GGAGTAAAAGAGCTGTATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACA | 867 | | | | | | |
| Db | 8462 | GGAGTAAAGAGAGCTGTATGAAACAATGGCAAGACCTCCAAGGAGAAATCGAAGCTCACACA | 8521 | | | | | | |

| | | | |
|----|------|--|------|
| Qy | 868 | GATGTTTATCACAAACCTGGATGAAACAGCCAAAAATCCTGAGATCCCTGGAAGGTTCC | 927 |
| Db | 8522 | GATATCTATCACAAACCTGGACGAAATGCCCCAAAAGTCTCTGAGATCCCTGGAAGGTTCT | 8581 |
| Qy | 928 | GATGATGCACTCCTGTTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCT | 987 |
| Db | 8582 | GACGATGACGCTTGTGTCAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGGCAACTT | 8641 |
| Qy | 988 | CGGAAAAAGTCTCTCAACATTAGGTCCCATTTTGGAAAGCCAGTTCTGACCAGTGGAAAGCGT | 1047 |
| Db | 8642 | CGGAAAAAGTCTCTCAACATTAGGTCTCACTTGGAAAGCCAGTTCTGACCAGTGGAAAGCGT | 8701 |
| Qy | 1048 | CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC | 1107 |
| Db | 8702 | CTGCACCTTTCTCTTCAGGAACCTTCTGGTATGGCTCCAGCTGAAAGATGATGATTAAGC | 8761 |
| Qy | 1108 | CGGCAGGCACTATTGGAGGCGACTTTCAGAGCAGTTTCAGAACAGAACGATGTACATAGG | 1167 |
| Db | 8762 | CGGCAGGCACTATTGGAGGAGACTTTCAGAGCGGTGCAGAACGAGAAATGATGTACACAGG | 8821 |
| Qy | 1168 | GCCTTCAAGAGGGGAATTGAAAACTTAAAGAACCTGTAAATCATAGTACTCTTGAGACTGTA | 1227 |
| Db | 8822 | GCCTTCAAGAGGGGAATTGAAAAACGAAAGAACCTGTAAATCATAGTACTCTTGAGACTGTA | 8881 |
| Qy | 1228 | CGAATATTTCTGACAGAGCAGCCCTTTTGGAAAGGACTTAGAGAAACTCTACAGGAGCCCCAGA | 1287 |
| Db | 8882 | CGAATATTTCTGACAGAGCAGCCCTTTTGGAAAGGACTTAGAGAAACTCTACAGGAGCCCCAGA | 8941 |
| Qy | 1288 | GAGCTGCCTCTGAGGAGAGAGCCCGCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG | 1347 |
| Db | 8942 | GAGCTGCCTCTGAGGAGAGAGCCCGCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG | 9001 |
| Qy | 1348 | GAGGTCAATACTGAGTGGGAAAAATTTGAACTTCCGCTGACTGGCAGAGAAAAATA | 1407 |
| Db | 9002 | GAGGTCAACACTCAGTGGGAAAAACTGAAAGCTGCACTCTGCAGACTGGCAGAGAAAAATA | 9061 |
| Qy | 1408 | GATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAG | 1467 |
| Db | 9062 | GACGAGGCCCTCGAAAGACTCCAGGAGCTTCAGGAAGCAACAGATGAGCTGGATCTCAAA | 9121 |
| Qy | 1468 | CTGCGCCAAGCTGAGGTGATCAAGGGATCTTGGCAGCCCGTGGCGGATCTCTCATTTGAC | 1527 |
| Db | 9122 | CTACGTCAAGGAGAGGTGATCAAGGATCTTGGCAGCCCTGTGGTGACCTCTCTCATTTGAC | 9181 |
| Qy | 1528 | TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGAGAAATTCGCGCTCTGAAA | 1587 |
| Db | 9182 | TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCCTTCGAGAGAAATTAACACCTCTGAAA | 9241 |
| Qy | 1588 | GAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACTTTTGGGCATTTCAGCTC | 1647 |
| Db | 9242 | GAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACTTTTGGGCATTTCAGCTC | 9301 |
| Qy | 1648 | TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAGGTG | 1707 |
| Db | 9302 | TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAGGTG | 9361 |
| Qy | 1708 | GCCGTGAGGACCGAGTCAGGACGTGATGAGCCCAAGCCCAAGGACTTTGGTCCAGCATCT | 1767 |
| Db | 9362 | GCCATTGAGGACCGCATCAGGACGTGATGAGCCCAAGGACTTTGGTCCAGCATCT | 9421 |
| Qy | 1768 | CAGCACCTTTCTTCCACGTCTGTCCAGGTCTTGGGAGAGGCCATCTCGCCAAACAAA | 1827 |
| Db | 9422 | CAGCACCTTTCTTCCACGTCTGTCCAGGTCTTGGGAGAGGCCATCTCGCCAAACAAA | 9481 |
| Qy | 1828 | GTGCCCTACTATATCAACACGAGACTCAAAACAACTTGTGGGAGAGGCCATCTCGCCAAACAAA | 1887 |
| Db | 9482 | GTGCCCTACTATATCAACACGAGACTCAAAACAACTTGTGGGAGAGGCCATCTCGCCAAACAAA | 9541 |
| Qy | 1888 | GAGCTCTACCCAGTCTTTAGCTGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCC | 1947 |
| Db | 9542 | GAGCTCTACCCAGTCTTTAGCTGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCC | 9601 |

| | | | |
|--|-----------------|---|---------------------|
| QY | 1948 | ATGAAACTCGAAGACTGCAGAAGGCGCTTTGCTTGGATCTCTTGAGCCTGTCAAGCTGCA | 2007 |
| Db | 9602 | ATGAAACTCGAAGACTGCAGAAGGCGCTTTGCTTGGATCTCTTGAGCCTATCGGCTGCA | 9661 |
| QY | 2008 | TGTGATGCTTGGACAGCACAACTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAG | 2067 |
| Db | 9662 | TGCGATGCTTGGACAGCACAACTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAG | 9721 |
| QY | 2068 | ATTATTAAATGTTTGACCACCTATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTC | 2127 |
| Db | 9722 | GTCATTAACTGTGACCACCTATTATGATCGCTAGAGCAAGAGCACAAATCTGGTC | 9781 |
| QY | 2128 | AACGTCCCTCTCTGCGTGGATATGTGTGAAGCTGGCTGGCTGAATGTTTATGATACGGGA | 2187 |
| Db | 9782 | AACGTCCCTCTCTGCGTGGATATGTGTCTCAATTTGGCTGCTGAATGTTTATGACACGGGA | 9841 |
| QY | 2188 | CGAACAGGGAGGATCCGTGTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA | 2247 |
| Db | 9842 | CGAACGGGGAGGATCCGGGTCTCTTTTAAACTGGCATCATTTCTCTGTGTAAAGCC | 9901 |
| QY | 2248 | CATTTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGT | 2307 |
| Db | 9902 | CATTTTGGAAGACAAGTACAGATACCTCTTCAAGCAAGTGGCAAGTTGACAGGATTTGT | 9961 |
| QY | 2308 | GACCAGCGAGCTGGCCCTCTCTGTCATGATTCTATCCAAATTTCAAGACAGTTGGGT | 2367 |
| Db | 9962 | GACCAGCGAGCTGGCCCTCTCTGTCATGACTCTATCCAGATCCCAAGACAGTTGGGT | 10021 |
| QY | 2368 | GAAGTTGCATCTTTGGGGGCAGTAACATTGAGC | 2401 |
| Db | 10022 | GAAGTCGCATCTGTCGGGGGCAGTAACATTGAGC | 10055 |
| RESULT 14 | | | |
| AX306153 | | | |
| LOCUS | | | |
| DEFINITION | | | |
| Sequence 904 from Patent WO0188188. | | | |
| AX306153 | | | |
| ACCESSION | | | |
| AX306153.1 GI:17645441 | | | |
| KEYWORDS | | | |
| Mus musculus (house mouse) | | | |
| SOURCE | | | |
| ORGANISM | | | |
| Mus musculus | | | |
| Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | |
| REFERENCE | | | |
| 1 | | | |
| Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y. | | | |
| Method for examining ischemic conditions | | | |
| Patent: WO 0188188-A 904 22-NOV-2001; | | | |
| School Juridical Person Nihon University (JP) | | | |
| FEATURES | | | |
| source | | | |
| 1..13815 | | | |
| /organism="Mus musculus" | | | |
| /mol_type="genomic DNA" | | | |
| /db_xref="taxon:10090" | | | |
| BASE COUNT | 4577 a | 2719 c | 3025 g 3494 t |
| ORIGIN | | | |
| Query Match | 62.8%; | Score 1506.8; | DB 6; Length 13815; |
| Best Local Similarity | 90.6%; | Pred. No. 0; | |
| Matches 1607; | Conservative 0; | Mismatches 167; | Indels 0; Gaps 0; |
| QY | 628 | TGGGCAACATCTGTAGATGGACAGAACCGCTGGGTCTTTTACAAAGACACTCATAGA | 687 |
| Db | 8207 | TGGGAAACATTCATAAAAGAGTAAGTAGCAAGAGGCTGCTTTGGAAGAACTCATAGA | 8266 |
| QY | 688 | TTACTGCAACAGTTCCCTGGACCTGGAAAGTTTCTTGCTGGCTTACAGAAGCTGAA | 747 |
| Db | 8267 | TTACTGCAGCAGTTCCCTCTGGACCTGGAGAGTTTCTTCTCGATTACGGAAGCAGAA | 8326 |
| QY | 748 | ACAACTGCCAATGTCTTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAG | 807 |
| Db | 8327 | ACAACTGCCAATGTCTTACAGGACGCTTCCCGTAAGGAGAAAGCTCCTAGAAGACTCCAGG | 8386 |

| | | | |
|----|------|---|------|
| QY | 808 | GGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAAGTGAAATTTGAAGCTCACACA | 867 |
| Db | 8387 | GGAGTCAGAGAGCTGATGAAACCCTATGGCAAGATCTCCAAGGAGAAAATTTGAAACTCACACA | 8446 |
| QY | 868 | GATGTTTTATCAAACTTGGATGAAACAGCCCAAAAATCCTGAGATCCCTGGAAAGTTCC | 927 |
| Db | 8447 | GATATCTATCAAACTTTGATGAAATTTGGCCCAAAAATCCTGAGATCCCTGGAAAGTTCC | 8506 |
| QY | 928 | GATGATGAGTCTCTGTTTACAAAGACGTTTGGATAACATGAATTTCAAGTGGAGTGAACCT | 987 |
| Db | 8507 | GATGAAGCACCCCTGTTTACAAAGACGTTTGGATAACATGAATTTCAAGTGGAGTGAACCT | 8566 |
| QY | 988 | CGGAAAAAGTCTCTCAACATTTAGTCCCATTTGGAAGCCAGTTCTGACCAAGTGGAAAGCGT | 1047 |
| Db | 8567 | CAGAAAAAGTCTCTCAACATTTAGTCCCATTTGGAAGCCAGTTCTGACCAAGTGGAAAGCGT | 8626 |
| QY | 1048 | CTGCACCTTTCTCTGAGGAACTTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC | 1107 |
| Db | 8627 | TTGCATCTTTCTCTGAGGAACTTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC | 8686 |
| QY | 1108 | CGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCCAGAGCAGAAACGATGTACATAGG | 1167 |
| Db | 8687 | CGTCAGGCACCCATCGTGGTGTATTTCCAGCAGTTTCCAGAGCAGAAATGATATACATAGG | 8746 |
| QY | 1168 | GCCTTCAAGAGGGAATTGAAAACCTAAAGAACCTGTAAATCATGAGTACTCTTTGAGACTGTA | 1227 |
| Db | 8747 | GCCTTCAAGAGGGAATTGAAAACCTAAAGAACCTGTAAATCATGAGTACTCTTTGAGACTGTA | 8806 |
| QY | 1228 | CGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACAGGAGCCCGAGA | 1287 |
| Db | 8807 | AGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACAGGAGCCCGAGA | 8866 |
| QY | 1288 | GAGCTGCCTCTCTGAGGAGAGAGCCCGAGAAATGTCACTCGGCTTTCTACGAAAGCAGGCTGAG | 1347 |
| Db | 8867 | GAACTGCCTCTCTGAAAGAAAGAGCTCAGAATGTCACTCGGCTTTCTACGAAAGCAGGCTGAA | 8926 |
| QY | 1348 | GAGTCAATACTGAGTGGGAAAAATTTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATA | 1407 |
| Db | 8927 | GAGTCAACCGCTGAATGGGACAAATTTGAACCTCGGCTCAGCTGATGGCAGAGAAAAATA | 8986 |
| QY | 1408 | GATGAGACCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAG | 1467 |
| Db | 8987 | GATGAAGCTCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAG | 9046 |
| QY | 1468 | CTGCGCCAAGCTGAGGTGATCAAGGGATCTCTGGCAGCCCGTGGCGCATCTCCTCATTTGAC | 1527 |
| Db | 9047 | TTGCGCCAAGCTGAGGTGATCAAGGGATCTCTGGCAGCCCGTGGCGCATCTCCTCATTTGAC | 9106 |
| QY | 1528 | TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTTCGAGGAGAAATTTGCGCCTCTGAAA | 1587 |
| Db | 9107 | TCTCTCCAAGATCACCTTGAAAGTCAAGGCACTTTCGAGGAGAAATTTGCGCCTCTGAAA | 9166 |
| QY | 1588 | GAGAACGTGAGCCACGTCAATGACCTTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTC | 1647 |
| Db | 9167 | GAGAATGTCAATCGTGTCAATGACCTTGCAATGACCTGACATCAGCTGACCACTGGGCATTCAGCTC | 9226 |
| QY | 1648 | TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTG | 1707 |
| Db | 9227 | TCACCTTATAACCTCAGCACTTTGGAAGATCTGAATACCAGATGGAGGCTTCTACAGGTG | 9286 |
| QY | 1708 | GCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGACCTTTGGTCCAGCATCT | 1767 |
| Db | 9287 | GCTGTGGAGGACCGTGTGACACAGCTGTCATGAAGCCCAAGGACCTTTGGTCCGTCATCC | 9346 |
| QY | 1768 | CAGCACTTTCTTTCCACGTCTGTCCAGGTTCCCTGGGAGAGAGCCCATCTGCCAAACAAA | 1827 |
| Db | 9347 | CAGCACTTTCTTTCCACGTCTGTCCAGGTTCCCTGGGAGAGAGCCCATCTGCCAAACAAA | 9406 |
| QY | 1828 | GTGCCCTACTATATCAACCCAGAGACTCAAAACAACTTGTCTGGGACCATCTCCAAAATGACA | 1887 |
| Db | 9407 | GTGCCCTACTATATCAACCCAGAGACCCAAACCACTTGTCTGGGACCATCTCCAAAATGACA | 9466 |
| QY | 1888 | GAGCTCTACCACTCTTAGCTGACCTGATAATGTGAGATTCTCAGCTTATAGGACTGCC | 1947 |

| | | | |
|---|------|---|------|
| Db | 9467 | GAGCTCTACCAAGTCTTTAGCTGACCTGAAATAATGTCAAGTTCTCCGGTATAGGACTGCC | 9526 |
| QY | 1948 | ATGAAGTCCGAAGACTGCAAGAGCCCTTTGGCTTGGATCTCTTGAGCCCTGTCAAGTACA | 2007 |
| Db | 9527 | ATGAAGCTCAGAAGGCTCCAGAAGAGCCCTTTGGCTTGGATCTCTTGAGCCCTGTCAAGTACA | 9586 |
| QY | 2008 | TGTGATGCTTGGACCAAGCAGCAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAG | 2067 |
| Db | 9587 | TGTGATGCTTGGACCAAGCAGCAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAG | 9646 |
| QY | 2068 | ATTATTAAATGTTTGAACCACTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTC | 2127 |
| Db | 9647 | ATAATTAAATGTTTGAACCACTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTC | 9706 |
| QY | 2128 | AACGTCCCTCTCTGCGTGGATATGCTGTGAACCTGGCTGCTGAATGTTTATGATACGGGA | 2187 |
| Db | 9707 | AATGTCCCTCTCTGCTGGATATGCTGTCAACTGGCTTCTCAATGTTTATGATACGGGA | 9766 |
| QY | 2188 | CGAACAGGAGGATCCGTGTCTCTCTTTTAAAACTGGCATCATTTCCCTGTGFAAAGCA | 2247 |
| Db | 9767 | CGAACAGGAGGATCCGTGTCTCTCTTTTAAAACTGGCATCATTTCTCTGTGFAAAGCA | 9826 |
| QY | 2248 | CATTGGAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGT | 2307 |
| Db | 9827 | CACCTGGAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACTGGCTTTGT | 9886 |
| QY | 2308 | GACCAGCGAGGCTGGCCCTCTCTGTCATGATTTCTATCCAAATTCGAAGACAGTTGGGT | 2367 |
| Db | 9887 | GACCAGCTAGGCTGGCTCTCTCTGTCATGATTTCTATCAAAATCCCAAGACAGTTGGGT | 9946 |
| QY | 2368 | GAAGTTGATCCTTTTGGGGGCGATTAACATTGAGC | 2401 |
| Db | 9947 | GAAGTTGCTCTTTTGGGGGCGATTAACATTGAGC | 9980 |
| RESULT 15 | | | |
| AX538582 | | | |
| LOCUS AX538582 13815 bp DNA linear PAT 23-NOV-2002 | | | |
| DEFINITION Sequence 2 from Patent WO0229056. | | | |
| ACCESSION AX538582 | | | |
| VERSION AX538582.1 GI:25271088 | | | |
| KEYWORDS Mus musculus (house mouse) | | | |
| SOURCE Mus musculus | | | |
| ORGANISM Mus musculus | | | |
| REFERENCE 1 | | | |
| AUTHORS Chamberlain, J.S. and Harper, S.Q. | | | |
| TITLE Mini-dystrophin nucleic acid and peptide sequences | | | |
| JOURNAL Patent: WO 0229056-A 2 11-APR-2002; | | | |
| THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US) | | | |
| FEATURES | | | |
| source Location/Qualifiers | | | |
| 1..13815 | | | |
| /organism="Mus musculus" | | | |
| /mol_type="genomic DNA" | | | |
| /db_xref="taxon:10090" | | | |
| BASE COUNT 4577 a 2719 c 3025 g 3494 t | | | |
| ORIGIN | | | |
| Query Match 62.8%; Score 1506.8; DB 6; Length 13815; | | | |
| Best Local Similarity 90.6%; Pred. No. 0; | | | |
| Matches 1607; Conservative 0; Mismatches 167; Indels 0; Gaps 0; | | | |
| QY | 628 | TGGGCAACATCTGTAGATGGACAGAACCCGCTGGGTTCTTTTACAAGACACTCATAGA | 687 |
| Db | 8207 | TGGGCAACATCTATAAAGAGTAAAGTACCAAGAGGCTGCTTTTGGAGAACTCATAGA | 8266 |
| QY | 688 | TTACTGCAACAGTTCCCTGGACCTGGAAAGTTTCTTGGCTGGCTTACAGAAGCTGAA | 747 |
| Db | 8267 | TTACTGCAACAGTTCCCTGGACCTGGAAAGTTTCTTCCCTGGATACGGAAGCAGAA | 8326 |
| QY | 748 | ACAACTGCCAATGTCTACAGGATGCTACCCGTAAGGAAAGGCTCTCTAGAAGACTCCAG | 807 |

| | | | |
|----|------|---|------|
| Db | 8327 | ACAACTGCCAATGTCTACAGGACGCTTCCCGTAAGGAGAGCTCTCTAGAAGACTCCAGG | 8386 |
| QY | 808 | GGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCCAAGGTGAAATTTGAAGCTCACACA | 867 |
| Db | 8387 | GGAGTCAAGAGAGCTGATGAAACAATGGCAAGATCTCCAAAGGAGAAATTTGAAGCTCACACA | 8446 |
| QY | 868 | GATGTTTATCACAACCTTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCC | 927 |
| Db | 8447 | GATATCTATCACAATCTTTGATGAAAAATGGCCAAAAAATCCTGAGATCCCTGGAAGGTTCC | 8506 |
| QY | 928 | GATGATGCAGTCTCTGTTTACAAAGACGTTTGGATTAACATGAATTTCAAGTGGAGTGAACCTT | 987 |
| Db | 8507 | GATGAAGCACCCCTGTTTACAAAGACGTTTGGATTAACATGAATTTCAAGTGGAGTGAACCTT | 8566 |
| QY | 988 | CGGAAAAAGTCTCTCAACATTAGGTCCCATTTTGGAAAGCCAGTTTCTGACCAGTGAAGCGT | 1047 |
| Db | 8567 | CAGAAAAAGTCTCTCAACATTAGGTCCCATTTTGGAAAGCCAGTTTCTGACCAGTGAAGCGT | 8626 |
| QY | 1048 | CTGCACTTTCTCTGCAAGGAATCTTGGTGTGGTGTGCTACAGCTGAAAGATGATGAATTAAGC | 1107 |
| Db | 8627 | TTGCATCTTTCTCTTCAAGGAATCTTGGTGTGGTGTGCTACAGCTGAAAGATGATGAATTAAGC | 8686 |
| QY | 1108 | CGGACGCACTTATTTGGAGGCACTTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGG | 1167 |
| Db | 8687 | CGTCAGGCACCCATCGGTGGTGAATTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGG | 8746 |
| QY | 1168 | GCCTTCAAGAGGGGAATTTGAAACTTAAAGAACCTGTAAATCATGAGTACTCTTCTGAGACTGTA | 1227 |
| Db | 8747 | GCCTTCAAGAGGGGAATTTGAAACTTAAAGAACCTGTAAATCATGAGTACTCTCTGAGACTGTA | 8806 |
| QY | 1228 | CGAATATTTCTGACAGAGCAGCCTTTTGGAAAGGACTTAGAGAAAACTCTACAGGAGCCCA | 1287 |
| Db | 8807 | AGAATATTTCTGACAGAGCAGCCTTTTGGAAAGGACTTAGAGAAAACTCTACAGGAGCCCA | 8866 |
| QY | 1288 | GAGCTGCTCTCTGAGGAGAGAGCCCAAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG | 1347 |
| Db | 8867 | GAAGTGCCTCTCTGAAGAAAGAGCTCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAA | 8926 |
| QY | 1348 | GAGTCAATACTGAGTGGGAAAAAATTTGAACCTTGCCTCGCTGACTGGCAGAGAAAAATA | 1407 |
| Db | 8927 | GAGTCAACGCTGAATGGGACAAATTTGAACCTTGCCTCGCTGACTGGCAGAGAAAAATA | 8986 |
| QY | 1408 | GATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGGCGCAGGATGAGCTGGACCTCAAG | 1467 |
| Db | 8987 | GATGAAGCTCTTGAAGACTCCAGGAACTTCAAGAGGCTGCGGATGAACCTGGACCTCAAG | 9046 |
| QY | 1468 | CTGGCCCAAGCTGAGGTGATCAAGGATCTTGGCAGCCCGTGGCGGATCTCTCATTTGAC | 1527 |
| Db | 9047 | TTGGCCCAAGCTGAGGTGATCAAGGATCTTGGCAGCCCGTGGCGGATCTCTCATTTGAC | 9106 |
| QY | 1528 | TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTTGCGCTCTGAAA | 1587 |
| Db | 9107 | TCTCTGCAAGATCACCTTGAAGAAAGTCAAGGCACCTTCGAGGAGAAATTTGACCTCTAAA | 9166 |
| QY | 1588 | GAGAACGTGAGCCACGTCATGACCTTGTCTGCGCCAGCTTACCACCTTTGGGCATTCAGCTC | 1647 |
| Db | 9167 | GAGAACGTGACCTGATGACCTTGTCTGCGCCAGCTTACCACCTTTGGGCATTCAGCTC | 9226 |
| QY | 1648 | TCACCGTATAACCTCAGCACTCTTGAAGACCTGAAACACCAAGATGGAAGCTTCTGCAGGTG | 1707 |
| Db | 9227 | TCACCGTATAACCTCAGCACTCTTGAAGATCTGAAATACCAAGATGGAGGCTTCTACAGGTG | 9286 |
| QY | 1708 | GCCGTGAGGACCGAGTCAAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG | 1767 |
| Db | 9287 | GCTGTGGAGGACCGTGTCAAGACAGCTGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG | 9346 |
| QY | 1768 | CAGCACTTTCTTTCCACGTCGTCTGTCAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG | 1827 |
| Db | 9347 | CAGCACTTTCTTTCCACGTCGTCTGTCAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG | 9406 |
| QY | 1828 | GTGCCCCTACTATATCAACCAAGAGCTCAACCAACTTGTGCGGAGGATGAGGATGAGGATGAGGATGAG | 1887 |

| | | | |
|----|------|---|------|
| Db | 9407 | GTGCCCTACTATATCAACACGAGACCCAAACACACTTGTGGGACCACCCCAAAATGACA | 9466 |
| Qy | 1888 | GAGCTCTACAGTCTTTAGCTGACCTGAATAATGTCAAGATTCTCAGCTTATAGGACTGCC | 1947 |
| Db | 9467 | GAGCTCTACAGTCTTTAGCTGACCTGAATAATGTCAAGTTCTCCGCGTATAGGACTGCC | 9526 |
| Qy | 1948 | ATGAAACTCCGAAGACTGCAGAAGGCCCTTGTCTTGGATCTCTTGAGCCTGTCAAGCTGCA | 2007 |
| Db | 9527 | ATGAAGCTCAGAAGGCTCCAGAAGGCCCTTGTCTTGGATCTCTTGAGCCTGTCAAGCTGCA | 9586 |
| Qy | 2008 | TGTGATGCCCTTGGACCAGCACAACTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAG | 2067 |
| Db | 9587 | TGTGATGCCCTTGGACCAGCACAACTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAG | 9646 |
| Qy | 2068 | ATTATTAAATGTTTGACCACACTATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTC | 2127 |
| Db | 9647 | ATAATTAACTGTTTGACTACAATTATGATCGTCTGGAGCAAGAGCACAAATCTGGTC | 9706 |
| Qy | 2128 | AACGTCCCTCTCTGCGTGGATATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGA | 2187 |
| Db | 9707 | AATGTCCCTCTCTGTGGATATGTGTCTCAACTGGCTTCTCAATGTTTATGATACGGGA | 9766 |
| Qy | 2188 | CGAACAGGAGGATCCGTGTCTCTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCA | 2247 |
| Db | 9767 | CGAACAGGAGGATCCGTGTCTCTCTTTTAAAACTGGCATCATTTCTCTGTGTAAAGCA | 9826 |
| Qy | 2248 | CATTGGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGT | 2307 |
| Db | 9827 | CATTGGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACTGGCTTTTGT | 9886 |
| Qy | 2308 | GACCAGCCGAGGCTGGGCCCTCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGT | 2367 |
| Db | 9887 | GACCAGCCGAGGCTGGGTCTTCTTCTGCATGATTCTATTCAAATCCCAAGACAGTTGGGT | 9946 |
| Qy | 2368 | GAAGTTGCATCCTTTGGGGGCGAGTAACATTGAGC | 2401 |
| Db | 9947 | GAAGTTGCTTCCTTTGGGGGCGAGTAACATTGAGC | 9980 |

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 11:04:35 ; Search time 3752.09 Seconds
(without alignments)
12961.661 Million cell updates/sec

Title: US-09-845-416-6_COPY_1000_3000
Perfect score: 2001
Sequence: 1 ggcagttcattgatggagag.....tcaaccacgagactcaaaca 2001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_estl:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rod:*
- 26: em_gss_phg:*
- 27: em_gss_vrl:*
- 28: gb_gss1:*
- 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 678.4 | 33.9 | 3870 | 11 BC036103 | BC036103 Homo sapi |
| 2 | 591.8 | 29.6 | 834 | 12 BI729851 | BI729851 603349511 |
| 3 | 547.4 | 27.4 | 3056 | 11 AK044536 | AK044536 Mus muscu |
| 4 | 491.4 | 24.6 | 750 | 12 BI730168 | BI730168 603349711 |

| | | | | | | |
|----|-------|------|------|----|----------|---------------------|
| 5 | 486.6 | 24.3 | 579 | 9 | AL121550 | AL121550 DKFZp762L |
| 6 | 471.6 | 23.6 | 728 | 14 | CB228986 | CB228986 AGENCOURT |
| 7 | 424.2 | 21.2 | 663 | 12 | BM488464 | BM488464 pgm2no.pk0 |
| 8 | 414.6 | 20.7 | 1490 | 11 | BC009242 | BC009242 Homo sapi |
| 9 | 409 | 20.4 | 595 | 14 | CB177816 | CB177816 is21c01.x |
| 10 | 389 | 19.4 | 644 | 13 | BU313510 | BU313510 603540290 |
| 11 | 324.6 | 16.2 | 402 | 14 | CB547284 | CB547284 AMGNNUC:S |
| 12 | 309.8 | 15.5 | 3753 | 11 | AK081426 | AK081426 Mus muscu |
| 13 | 265.4 | 13.3 | 3051 | 11 | BC036095 | BC036095 Homo sapi |
| 14 | 261.2 | 13.1 | 2334 | 11 | BC011062 | BC011062 Mus muscu |
| 15 | 257.8 | 12.9 | 1541 | 11 | AK034383 | AK034383 Mus muscu |
| 16 | 252.4 | 12.6 | 772 | 12 | BI250598 | BI250598 602993659 |
| 17 | 233 | 11.6 | 835 | 12 | BI553820 | BI553820 603190772 |
| 18 | 230.8 | 11.5 | 784 | 10 | BG212445 | BG212445 RST32032 |
| 19 | 230.2 | 11.5 | 763 | 14 | CB518960 | CB518960 UI-M-GH0- |
| 20 | 229.2 | 11.5 | 778 | 14 | CB524596 | CB524596 UI-M-FY0- |
| 21 | 226 | 11.3 | 250 | 10 | BF963618 | BF963618 QV2-NN004 |
| 22 | 216.6 | 10.8 | 581 | 13 | BU301653 | BU301653 603609005 |
| 23 | 199.8 | 10.0 | 696 | 12 | BJ075057 | BJ075057 BJ075057 |
| 24 | 199.2 | 10.0 | 697 | 13 | BU456556 | BU456556 603771141 |
| 25 | 182.4 | 9.1 | 645 | 9 | AW467977 | AW467977 he29g08.x |
| 26 | 179.6 | 9.0 | 466 | 14 | CD549993 | CD549993 B0305E01- |
| 27 | 172.2 | 8.6 | 642 | 9 | AL855376 | AL855376 AL855376 |
| 28 | 167.6 | 8.4 | 502 | 9 | AL602076 | AL602076 DKFZp313B |
| 29 | 167.2 | 8.4 | 646 | 12 | BI289102 | BI289102 UI-R-DK0- |
| 30 | 165.4 | 8.3 | 851 | 13 | BU201022 | BU201022 603952191 |
| 31 | 164.6 | 8.2 | 600 | 12 | BI988528 | BI988528 4012-24 M |
| 32 | 158 | 7.9 | 405 | 14 | CB811308 | CB811308 AMGNNUC:S |
| 33 | 153.4 | 7.7 | 481 | 10 | BE557463 | BE557463 f103h06.Y |
| 34 | 153.2 | 7.7 | 885 | 11 | AK020881 | AK020881 Mus muscu |
| 35 | 150.6 | 7.5 | 515 | 28 | AZ780914 | AZ780914 2M0018010 |
| 36 | 144.6 | 7.2 | 675 | 14 | BY706879 | BY706879 BY706879 |
| 37 | 144.2 | 7.2 | 392 | 10 | AW948794 | AW948794 QV0-FT000 |
| 38 | 142.2 | 7.1 | 426 | 13 | BQ319056 | BQ319056 IL5-CT051 |
| 39 | 141.6 | 7.1 | 794 | 28 | BZ097143 | BZ097143 CH230-236 |
| 40 | 141 | 7.0 | 761 | 28 | BZ129893 | BZ129893 CH230-452 |
| 41 | 140.6 | 7.0 | 649 | 13 | BQ209933 | BQ209933 UI-R-EP0- |
| 42 | 140 | 7.0 | 515 | 10 | BB651473 | BB651473 BB651473 |
| 43 | 137.8 | 6.9 | 479 | 13 | BQ304046 | BQ304046 QV2-BT063 |
| 44 | 134 | 6.7 | 828 | 28 | BZ126085 | BZ126085 CH230-452 |
| 45 | 133.2 | 6.7 | 521 | 28 | AZ226551 | AZ226551 RPCI-23-9 |

ALIGNMENTS

| | | | | | |
|------------|--|---------|------|--------|-----------------|
| RESULT 1 | BC036103 | 3870 bp | mRNA | linear | HTC 04-MAR-2003 |
| LOCUS | BC036103 Homo sapiens, clone IMAGE:5274415, mRNA. | | | | |
| DEFINITION | BC036103 | | | | |
| ACCESSION | BC036103.1 GI:23271310 | | | | |
| VERSION | HTC. | | | | |
| KEYWORDS | Homo sapiens (human) | | | | |
| SOURCE | Homo sapiens | | | | |
| ORGANISM | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | |
| REFERENCE | 1 (bases 1 to 3870) | | | | |
| AUTHORS | Strausberg,R. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA | | | | |
| REMARK | NIH-MGC Project URL: http://mgc.nci.nih.gov | | | | |
| COMMENT | Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN) cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 | | | | |

Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 48 Row: f Column: 7
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein
This clone has the following problem: retained intron.

FEATURES
source
1. .3870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5274415"
/tissue_type="Brain, hippocampus"
/clone_lib="NIH_MGC_95"
/lab_host="DH10B"
/note="Vector: pBluescript"

BASE COUNT 1321 a 677 c 763 g 1109 t
ORIGIN

Query Match 33.9%; Score 678.4; DB 11; Length 3870;
Best Local Similarity 97.7%; Pred. No. 1.9e-126;
Matches 688; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAA 60
Db 1143 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAA 1202
QY 61 GTATTATCGGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAAGGAGATTTCTAAT 120
Db 1203 GTATTATCGGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAAGGAGATTTCTAAT 1262
QY 121 GATGTGGAAGTGGTGAAGACCAGTTTCATACTCATGAGGGTACATGATGGATTGACA 180
Db 1263 GATGTGGAAGTGGTGAAGACCAGTTTCATACTCATGAGGGTACATGATGGATTGACA 1322
QY 181 GCCCATCAGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240
Db 1323 GCCCATCAGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 1382
QY 241 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 300
Db 1383 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 1442
QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTACATAGAGTTTAAATG 360
Db 1443 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTACATAGAGTTTAAATG 1502
QY 361 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGA 420
Db 1503 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGA 1562
QY 421 ACAAGGAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTA 480
Db 1563 ACAAGGAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTA 1622
QY 481 CAACAACATAAGTGCTTCAAGAAGATCTAGAACAAGAACAAGTCAAGGTCAATTCTCTC 540
Db 1623 CAACAACATAAGTGCTTCAAGAAGATCTAGAACAAGAACAAGTCAAGGTCAATTCTCTC 1682
QY 541 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAGCAAACTGCTGCTTTGGAA 600
Db 1683 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAGCAAACTGCTGCTTTGGAA 1742
QY 601 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCGC 660
Db 1743 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCGC 1802

QY 661 TGGGTTCTTTTACAAGACCAGCTGACCTAGCTAGTCTCTGGACTGAC 704
Db 1803 TGGGTTCTTTTACAAGACATCCTTCTCAAATGGCAACGCTCTTAC 1846

RESULT 2
BI729851 834 bp mRNA linear EST 20-SEP-2001
LOCUS 603349511P1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5357162 5',
DEFINITION mRNA sequence.
ACCESSION BI729851
VERSION BI729851.1 GI:15706864
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 834)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11908 row: e column: 03
High quality sequence stop: 796.

FEATURES
source
1. .834
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:5357162"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_94"
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 275 a 173 c 199 g 187 t
ORIGIN

Query Match 29.6%; Score 591.8; DB 12; Length 834;
Best Local Similarity 90.5%; Pred. No. 4.7e-109;
Matches 665; Conservative 0; Mismatches 67; Indels 3; Gaps 3;

QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTCCTG 872
Db 75 GGAAGAACTCATAGATTACTGCAGCAGTTCCCTCTGGACCTGGAGAAGTTTCTTCCTG 134
QY 873 GCTTACAGAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTACCCGTAAGGAAAGCT 932
Db 135 GATTACGGAAGCAGAAACAACCTGCCAATGTCTTACAGGACGCTTCCCGTAAGGAGAAGCT 194
QY 933 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAAGTGA 992
Db 195 CCTAGAAGACTCC-AGGGAGTCAGAGAGCTGATGAAACCATGGCAAGATCTCCAAAGAGA 253
QY 993 AATTGAAGCTCACACAGATGTTTATCAACACCTGGATGAAACAGCCAAAAATCCTGAG 1052
Db 254 AATTGAAGCTCACACAGATATCTATCACAATCTTGTATGAAAAATGGCCAAAAATCCTGAG 313
QY 1053 ATCCCTGGAAGGTTCCGATGATGAGTCCCTGTTTACAAAGACGTTTGGATAACATGAATT 1112
Db 314 ATCCCTGGAAGGTTCCGATGATGAGCAGCCCTGTTTACAAAGACGTTTGGATAACATGAATT 373

2Y 1113 CAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTGTGGAAGCCAGTTC 1172
|||||
2b 374 CAAGTGGAGTGAACCTTCAGAAAAAGTCTCTCAACATTAGGTCCCATTGTGGAAGCAAGTTC 433
|||||
2Y 1173 TGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGTGGTGTTGCTACAGCTGAA 1232
|||||
2b 434 TGACCAGTGGAAAGCGTTTGGCATCTTTCTCTTCAGGAACTTCTTGTGTTGGCTACAGCTGAA 493
|||||
2Y 1233 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCGCACTTCCAGCAGTTCAGAAAGCA 1292
|||||
2b 494 AGATGATGAACCTGAGCCGTCAGGCACCCCATCGGTGGTGTAATTTCCAGCAGTTCAGAAAGCA 553
|||||
2Y 1293 GAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG 1352
|||||
2b 554 GAATGATATACATAGGCGCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG 613
|||||
2Y 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTTGGAAAGGACTAGAGAAACT 1412
|||||
2b 614 TACTCTGGAGACTGTGAGATATTTCTGACAGAGCAGCCCTTTTGGAAAGGACTAGAGAAACT 673
|||||
2Y 1413 C-TACCAGGAGCCAGAGAGTGCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTC 1471
|||||
2b 674 CTTACCAGGAGCCAGAGAACTGCCTCCTGAGAAAGAGCTCAGAAATGTCACTCGGCTCC 733
|||||
2Y 1472 TACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACTTGCACCTC-CGCT 1530
|||||
2b 734 TACGAAAGCAGGCTGAGAGGTCAACGCTGAATGGGACCAATTGAACTTCGGCTCAAGCT 793
|||||
Y 1531 GACTGGCAGAGAAAA 1545
|||||
b 794 GATTGGCAGAGAAAA 808
|||||
RESULT 3
AK044536 3056 bp mRNA linear HTC 05-DEC-2002
LOCUS Mus musculus adult retina cDNA, RIKEN full-length enriched library,
DEFINITION clone:A930019F21 product:dystrophin, muscular dystrophy, full
insert sequence.
ACCESSION AK044536
VERSION AK044536.1 GI:26090404
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE Carninci,P. and Hayashizaki,Y.
AUTHORS High-efficiency full-length cDNA cloning
TITLE Meth. Enzymol. 303, 19-44 (1999)
JOURNAL 99279253
MEDLINE 10349636
PUBMED 10349636
2
REFERENCE Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
AUTHORS Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
JOURNAL prepare full-length cDNA libraries for rapid discovery of new genes
MEDLINE Genome Res. 10 (10), 1617-1630 (2000)
PUBMED 20499374
REFERENCE 11042159
3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitesunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yanamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861

4
REFERENCE Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
AUTHORS Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
and Hayashizaki,Y.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409 (6821), 685-690 (2001)
MEDLINE 21085660
PUBMED 11217851
5
REFERENCE The PANTOM Consortium and the RIKEN Genome Exploration Research
AUTHORS Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
JOURNAL of 50,770 full-length cDNAs
REFERENCE Nature 420, 563-573 (2002)
AUTHORS 6 (bases 1 to 3056)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Katoch,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gs.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Retina RNA was provided by Dr. Stefano Gustincich (Department of
Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA
02115, USA) whose assistance is gratefully acknowledged. Please
visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/
Location/Qualifiers
1. .3056
/organism="Mus musculus"
/mol_type="mrna"
/strain="C57BL/6J"
/db_xref="FANTOM DB:A930019F21"
/db_xref="taxon:10090"
/clone="A930019F21"
/tissue_type="retina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"

[illegible][illegible]

| | | | |
|----|------|---|------|
| QY | 1217 | TGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCGCACTTTC | 1276 |
| Db | 311 | TGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCGCACTTTC | 370 |
| QY | 1277 | CAGCAGTTCAGAAGCAGAAACGATGTACATAGGGCCTTCAAGAGGGGAATTGAAAACTAAAG | 1336 |
| Db | 371 | CAGCAGTTCAGAAGCAGAAACGATGTACATAGGGCCTTCAAGAGGGGAATTGAAAACTAAAG | 430 |
| QY | 1337 | AACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGG | 1396 |
| Db | 431 | AACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGG | 490 |
| QY | 1397 | AAGGACTAGAGAAAACCTTACCAGGAGCCCAAGAGAGCTGCCTCCTGAGGAGAGAGCCCCAGA | 1456 |
| Db | 491 | AAGGGCTAGAGAAAACCTTACCAGGAGCCCAAGAGAGCTGCCTCCTGAGGAGAGAGCCCCAG | 550 |
| QY | 1457 | ATGTCACCTGGCCTTCTACGAAAGCAGGCT | 1485 |
| Db | 551 | AATGTCACCTGGCCTTCTACGAAAGCAGGCT | 579 |

| | |
|------------|---|
| RESULT 6 | |
| CB228986 | |
| LOCUS | 728 bp mRNA linear EST 10-FEB-2003 |
| DEFINITION | AGENCOURT_11499247 NICHD_Rh_Ov1 Macaca mulatta cDNA clone |
| | IMAGE:6884820 5', mRNA sequence. |
| ACCESSION | CB228986 |
| VERSION | CB228986.1 GI:28280564 |
| KEYWORDS | EST. |
| SOURCE | Macaca mulatta (rhesus monkey) |
| ORGANISM | Macaca mulatta |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Cercopithecinae; Macaca. |
| REFERENCE | 1 (bases 1 to 728) |
| AUTHORS | NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap . |
| TITLE | National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index |
| JOURNAL | Unpublished |
| COMMENT | Contact: Robert Strausberg, Ph.D. |

```

BASE COUNT      256 a      127 c      162 g      182 t      1 others
ORIGIN

Query Match      23.6%;      Score 471.6;      DB 14;      Length 728;
Best Local Similarity 95.3%;      Pred. No. 8.1e-85;
Matches 486;      Conservative      0;      Mismatches 24;      Indels 0;      Gaps 0;

QY      195      GGTGGTAATTCTACAATTGGGAAGTAAGCTGATTGGACACGAAAAATTATCAGAAGA      254

```

Db 1 GGTGGTAATATTCTACAATTGGGAAGTCAGCTGATTGGAACAGGGAAATTATCAGAAGA 60

QY 255 TGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTTAATTAATCAAGATGGGAATGCCTCAG 314

Db 61 TGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTTAATTAATCAAGATGGGAATGCCTCAG 120

QY 315 GGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCA 374

Db 121 GGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCA 180

QY 375 GAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAAGAAAGAAACAAGGAAATGGA 434

Db 181 GAAACTGAAAGAGTTGAATGACTGGCTGACAAAAACAAGAAAGAAACAAGGAAATGGA 240

QY 435 GGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTACAAACAATAAGGT 494

Db 241 GAAAGAACCCCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTACAAACAATAAGGT 300

QY 495 GCTTCAAGAGATCTAGAACAAAGAACAGTCAGGGTCAATCTCTCACTCACAATGAGGT 554

Db 301 GCTTCAAGAGATCTAGAACAAAGAACAGTCAGGGTCAATCTCTCACTCACAATGAGGT 360

QY 555 GGTAGTTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAACAACTTAAGGT 614

Db 361 GGTAGTTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAGCAACTTAAGGT 420

QY 615 ATTGGGAGATCGATGGCAAAACATCTGATGAGATCAGCAACTGCTGCTTTGGAAGAACAACTTAAGGT 674

Db 421 ATTGGGAGATCGATGGCAAAACATCTGATGAGATGAGCAAGACCGCTGGGTTCTTTTACA 480

QY 675 AGACCAAGCCTGACCTAGCTCCTGGAGTGCAC 704

Db 481 AGACATCCTTCTCAAAATGGCAACGCTCTAC 510

RESULT 7

BM488464

LOCUS

DEFINITION

BM488464.1 GI:18609395 663 bp mRNA linear EST 07-FEB-2002

gpm2n.pk007.117 Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library (pgm2n) Gallus gallus CDNA clone pgm2n.pk007.117 5' similar to sp|P11533|DMD CHICK Dystrophin pir|S02041|dystrophin, muscle - chicken emb|CAA31746.1| (X13369) dystrophin (AA 1 - 3660) [Gallus gallus], mRNA sequence.

ACCESSION

BM488464

VERSION

KEYWORDS

SOURCE

ORGANISM

Gallus gallus (chicken)

Gallus gallus

Fukayota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

1 (bases 1 to 663)

Cogburn, L.A. and Monsonego-Ornan, E.

ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library, USDA/IFAPS Animal Genome Project

JOURNAL

Unpublished

CONTACT: Larry A. Cogburn

University of Delaware

Townsend Hall, Newark, DE 19717, USA

Tel: 302-831-1335

Fax: 302-831-2822

Email: cogburn@udel.edu, www.chickest.udel.edu.

FEATURES

source

1..663

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="Commercial broiler and Ottawa Res. Centre strains 90 & 21"

/db_xref="taxon:9031"

/clone="pgm2n.pk007.117"

/sex="Male and Female"

/tissue_type="Breast muscle, leg muscle and epiphyseal

growth plate"

/dev_stage="Breast, leg:Embryo(d19); post-hatch(1d,1,3,5,7,9,11 weeks); growth plate(1d,7d,14d post-hatch)"

/lab_host="E. coli EMDH10B"

/clone_lib="Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library (pgm2n)"

/note="Vector: pCMVSPORT6; Library made from equivalent pools of total RNA isolated from each tissue (embryonic muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth plate 33.3% of the final RNA pool). Single pass sequencing from 5'-end"

BASE COUNT 209 a 132 c 172 g 144 t 6 others

ORIGIN

Query Match 21.2%; Score 424.2; DB 12; Length 663;

Best Local Similarity 78.1%; Pred. No. 2.9e-75;

Matches 518; Conservative 0; Mismatches 144; Indels 1; Gaps 1;

QY 960 GCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCA 1019

Db 1 GCTCATGAAGCAGTGGCAGGATCTACAGGCAGAAATTGATGCACATCT-TNN 59

QY 1020 CAACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAAGTTCCGATGATGTCAGT 1079

Db 60 CAACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAAGTTCCGATGATGTCAGT 119

QY 1080 CCTGTTACAAAGACGTTTGGATTAACATGAACCTTCAAGTGGAGTGAACTTCGGAAAAAGTC 1139

Db 120 CCTGTTGCAGAGACGTCGTGATTAACATGAACCTTCAAGTGGAGTGAGCTTAGGAAGAAATC 179

QY 1140 TCTCAACATTAGTCTCCATTGGAAAGCCAGTTCTGACAGTGGAAAGCGTCTGCACCTTC 1199

Db 180 TCTAAACATTAGATCTCATTGGAAAGCCAGCACAGACCAGTGGAAAGCGTTTACATCTCTC 239

QY 1200 TCTCAGGAACCTTCTGGTGTGGTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACC 1259

Db 240 TCTTCAGGAACCTTCTGGCATGGTGTGCAATTGAAGGAGGATGAATTAACACAGCAAGCAC 299

QY 1260 TATTGGAGGCGACTTTCAGCAGTTCCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAG 1319

Db 300 CATTGGTGGAGATATTCACACTGTGAGAAAGCAGAAATGATGTTTCATAGGACTTTCAGAG 359

QY 1320 GGAATTGAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCT 1379

Db 360 GGAGCTGAAACAAAGAACCTGTTATCATGAATGCACCTTGAGACTGTGCGACTCTTCTCT 419

QY 1380 GACAGAGCAGCCTTTTGAAGGACTAGAGAAACTCTACCAGAGCCAGAGAGCTGCCTCC 1439

Db 420 GGCAGATCAACAGTAGAGGAGCTGGAAAGGTTCTATCCAGAACCAAGAGACCTATCACC 479

QY 1440 TGAGGAGAGAGCCAGAGATGTCACTCGGCTTCTACGAAAGCAGAGGCTGAGGAGGTCAATAC 1499

Db 480 TGAGGAGAGAGCCAGAGATGTCACTAAAGTTCTCCGAGGCAAGCAGATGATGTCAGNAC 539

QY 1500 TGAGTGGGAAAAATTTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCT 1559

Db 540 TGAGTGGGATAAGCTAAATCTACGTTCTGCTGATTGGCAAAAGAGATAGATGATGCTCT 599

QY 1560 TGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGGCCAAGC 1619

Db 600 TGAAGACTGCAGGGTCTTCCAGGAGGCAATGGATGAACCTNNCTTGAAACTGCGCCAGGC 659

QY 1620 TGA 1622

Db 660 TGA 662

RESULT 8

BC009242

LOCUS

DEFINITION

BC009242 1490 bp mRNA linear HTC 04-MAR-2003

Homo sapiens, dystrophin (muscular dystrophy, Duchenne and Becker types), includes DXS142, DXS164, DXS206, DXS230, DXS239, DXS268, DXS269, DXS270, DXS272, clone IMAGE:3029414, mRNA.

ACCESSION BC009242
VERSION BC009242.1 GI:14714379
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1490)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (06-JUN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbiology.org>
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 10 Row: j Column: 10
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 5032284
This clone has the following problem: retained intron.
FEATURES
source
1..1490
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3029414"
/tissue_type="Muscle, rhabdomyosarcoma"
/clone_lib="NIH MGC 17"
/lab_host="DH10B-R"
/note="Vector: pOTB7" 358 t
BASE COUNT 505 a 299 c 328 g 358 t
ORIGIN
Query Match 20.7%; Score 414.6; DB 11; Length 1490;
Best Local Similarity 99.0%; Pred. No. 2.8e-73;
Matches 417; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAA 60
DB 1069 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAA 1128
QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTCTTAAT 120
DB 1129 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTCTTAAT 1188
QY 121 GATGTGGAAGTGGTGAAGAACCCAGTTTCATACTCATGAGGGGTACATGATGGATTGACA 180
DB 1189 GATGTGGAAGTGGTGAAGAACCCAGTTTCATACTCATGAGGGGTACATGATGGATTGACA 1248
QY 181 GCCCATCAGGCCCGGTTGGTAATATTTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240
DB 1249 GCCCATCAGGCCCGGTTGGTAATATTTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 1308
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAAGACAGATGAATCTCCTAAATTCAAGA 300
DB 1309 AAATTATCAGAAGATGAAGAACTGAAGTACAAAGACAGATGAATCTCCTAAATTCAAGA 1368
QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATG 360
DB 1369 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATG 1428

QY 361 GATCTCCAGATCAGAAACTGAAGAGATTGAATGACTGGCTAACAAAAACAGAGAAGA 420
DB 1429 GATCTCCAGATCAGAAACTGAAGAGATTGAATGACTGGCTAACAAAAACAGAGAAGA 1488
QY 421 A 421
DB 1489 A 1489
RESULT 9
CB177816/c
LOCUS CB177816 595 bp mRNA linear EST 31-JAN-2003
DEFINITION is21c01.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6553129 3'
similar to SW:DMD_HUMAN P11532 DYSTROPHIN. [1] ; mRNA sequence.
ACCESSION CB177816
VERSION CB177816.1 GI:28186206
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 595)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
M., Gibbons,M., McCann,R., Cole,R., Tsagarishvili,R., Williams,T.,
Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished
TITLE Other_ESTs: is21c01.y1
JOURNAL Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
COMMENT Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 448.
FEATURES
source
Location/Qualifiers
1..595
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6553129"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_lib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
Size-selected on agarose gel. Average insert size ~1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."
BASE COUNT 146 a 136 c 116 g 197 t
ORIGIN
Query Match 20.4%; Score 409; DB 14; Length 595;
Best Local Similarity 100.0%; Pred. No. 3.4e-72;
Matches 409; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAA 60
DB 409 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAA 350

| | | | |
|----|-----|---|-----|
| QY | 61 | GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTTAAT | 120 |
| | | | |
| DB | 349 | GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTTAAT | 290 |
| | | | |
| QY | 121 | GATGTGGAAGTGGTGAAGACCAGCTTTCATACTCATGAGGGGTACATGATGGATTGACA | 180 |
| | | | |
| DB | 289 | GATGTGGAAGTGGTGAAGACCAGCTTTCATACTCATGAGGGGTACATGATGGATTGACA | 230 |
| | | | |
| QY | 181 | GCCCATCAGGGCCGGGTGGTAAATATCTACAATTGGGAAGTAAGCTGATTGGAACAGGA | 240 |
| | | | |
| DB | 229 | GCCCATCAGGGCCGGGTGGTAAATATCTACAATTGGGAAGTAAGCTGATTGGAACAGGA | 170 |
| | | | |
| QY | 241 | AAATTATCAGAAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGA | 300 |
| | | | |
| DB | 169 | AAATTATCAGAAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGA | 110 |
| | | | |
| QY | 301 | TGGGAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAATG | 360 |
| | | | |
| DB | 109 | TGGGAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAATG | 50 |
| | | | |
| QY | 361 | GATCTCCAGAATCAGAACTGAAGAGTTGAAATGACTGGCTAACAAAA | 409 |
| | | | |
| DB | 49 | GATCTCAGAAATCAGAACTGAAGAGTTGAAATGACTGGCTAACAAAA | 1 |

| | | | | | |
|------------|--|-------------|---------------|------------|----------------------|
| RESULT 10 | BU313510 | 644 bp | mRNA | linear | EST 28-NOV-2002 |
| LOCUS | 603540290F1 | CSEQCHN61 | Gallus gallus | CDNA clone | ChEST508c24 5', mRNA |
| DEFINITION | sequence. | | | | |
| ACCESSION | BU313510 | | | | |
| VERSION | BU313510.1 | GI:25821511 | | | |
| KEYWORDS | EST. | | | | |
| SOURCE | Gallus gallus (chicken) | | | | |
| ORGANISM | Gallus gallus | | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| | Archosauria; Aves; Neognathae; Galliformes; Phasianidae; | | | | |
| | Phasianinae; Gallus. | | | | |
| | 1 (bases 1 to 644) | | | | |
| REFERENCE | Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., | | | | |
| AUTHORS | Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J. | | | | |
| | A Comprehensive Collection of Chicken cDNAs | | | | |
| TITLE | Curr. Biol. 12 (22), 1965-1969 (2002) | | | | |
| JOURNAL | | | | | |

compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

| BASE COUNT | 192 a | 141 c | 168 g | 142 t | 1 others |
|------------|-------|-------|-------|-------|----------|
| ORIGIN | | | | | |

| | | | | |
|-----------------------|--------------|--------------------|-----------------|-------------|
| Query Match | 19.4%; | Score 389; | DB 13; | Length 644; |
| Best Local Similarity | 75.5%; | Pred. No. 3.7e-68; | | |
| Matches 482; | Conservative | 0; | Mismatches 156; | Indels |

| | | | | | | | | | | | | | | | | |
|----|------|-------------|----------|------------|----------|----------|---------|----------|-----------|-----------|------------|---------|------|--------|-------|-----|
| QY | 1166 | CCAGTTCTGAC | CAGTGGAA | CGCTTCTCTG | CAGGAAC | TTCTCTG | CGAGCA | CTTCTGGT | GTGGCTAC | 1222 | | | | | | |
| DB | 7 | CGAGC | CAGAC | CAGTGGAA | CGGTTTAC | ATCTCTCT | TCAGGA | ACTTTTGG | CATGGCTGC | 66 | | | | | | |
| QY | 1226 | AGCTGAA | AGATGAT | GAATTAAG | CCGCG | CAGGCAC | CTATTGG | AGGCG | ACTTTCCAG | CAGTTTC | 1285 | | | | | |
| DB | 67 | AATTGA | AGGAGG | ATGAATTA | AAAA | CAGCAAG | CACCCAT | TGTTGG | AGATATCC | CACTGTGC | 126 | | | | | |
| QY | 1286 | AGAAG | CAGAA | CGATG | TACATAG | GGCCTT | CAAGAG | GGAATTGA | AAAA | CTAAAGAA | CCTGTAA | 1345 | | | | |
| DB | 127 | AGAAG | CAGAA | TGATG | TTTATAG | GACTTT | CAAGAG | GAGCTG | AAAA | CAAAAAGAA | CCTGTTA | 186 | | | | |
| QY | 1346 | TCATG | AGTACT | CTTTG | AGACTGT | ACGAAT | ATTCTG | ACAG | AGCAG | CCCTTTG | GAAGGACTAG | 1405 | | | | |
| DB | 187 | TCATG | AATG | CACTTG | AGACTGT | GCGACT | TTCTCT | GGCAG | ATCA | ACCAGT | AGAGGACTGG | 246 | | | | |
| QY | 1406 | AGAA | ACTCTAC | CAGGAG | CCCGAG | AGCTGC | CTCTCT | GAGG | AGAGAG | CCCGAG | AATGTCACTC | 1465 | | | | |
| DB | 247 | AAA | AGTCTAT | CCAGAA | CCCAAG | AGACCT | ATCAC | CTGAG | GAGAG | GCGCCAG | AATGTCACTA | 306 | | | | |
| QY | 1466 | GGCTT | CTAC | GAAAG | CAGGCTG | AGGAGGT | CAATA | CTAGT | GTGG | AAAAAATTG | AACCTGC | 1525 | | | | |
| DB | 307 | AAGTT | CTCC | GAAAG | CAAC | GATGAT | GTCA | GAAC | TGAGT | GGGATA | AAGCTAA | CTACGTT | 366 | | | |
| QY | 1526 | CCGCT | GACTGG | CAGAG | AAAAAT | AGATG | AGAC | CCCTTG | AAAG | ACTCCAG | GAACTTCA | AGAGG | 1585 | | | |
| DB | 367 | CTGCT | GATGG | CAAA | AGAAG | ATAG | ATGAT | GTCTCT | TGAA | AGACTG | CAGGGT | CTTCAG | GAGG | 426 | | |
| QY | 1586 | CCACG | GATG | AGCTGG | ACCTCA | AGCTG | CGCC | CAAGCTG | AGGTG | ATCA | AGGGAT | CCTGG | CAGC | 1645 | | |
| DB | 427 | CAATG | GATGA | ACTAG | ACCTGA | AACTG | CGCC | AGGCTG | AAG | CAATCA | AGGGAT | CCTGG | CAGC | 486 | | |
| QY | 1646 | CGTGG | CGATCT | CCTCAT | TGACTCT | CTCCA | AGATCA | CCCTCG | AGAA | AGTCA | AGGC | ACTTC | 1705 | | | |
| DB | 487 | CAGTGG | GGATCT | GCTG | ATAG | ACTCTCT | CG | AGGAT | CACTTAG | AAAAAGT | CAAGTTT | ATC | 546 | | | |
| QY | 1706 | GAGG | AGAAAT | TGCG | CCTCTG | AAAG | AGAA | CGTG | AGCC | ACGTCA | ATGAC | CTTGCT | CGCC | AGC | 1765 | |
| DB | 547 | GAGC | AGAAAT | TG | TGCG | CCCTTTAN | AG | AGAA | GGTGC | ATCA | AGTCA | ATG | AGCT | GGCTCA | CCGGT | 606 |
| QY | 1766 | TTACC | ACTTTG | GGCAT | TT | CAGCT | CT | CA | CCGT | ATA | ACCTC | 1803 | | | | |
| DB | 607 | TCG | TCC | CCCTG | ATAT | TCAG | TTCT | CC | CCAT | ACACT | CTC | 644 | | | | |

RESULT 11
CB547284/C

| | | | | |
|------------|---|------|--------|-----------------|
| CB547284 | 402 bp | mRNA | linear | EST 01-APR-2003 |
| LOCUS | CB547284 | | | |
| DEFINITION | AMGNNUC:SRPB2-00242-E3-A srpb2 (10220) Rattus norvegicus CDNA clone | | | |
| | srpb2-00242-e3 5', mRNA sequence. | | | |

ACCESSION
CB547284

VERSION CB547284.1 GI:29431225

KEYWORDS EST.

SOURCE *Rattus norvegicus* (Norway rat)

ORGANISM

Eukaryota: Metazoa

Mammalia: Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE 1 (bases 1 to 402)

REFERENCE

THE UNIVERSITY OF CHICAGO

Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>.

FEATURES

Location/Qualifiers
1. 3753
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:C130016K19"
/db_xref="taxon:10090"
/clone="C130016K19"
/tissue type="head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days embryo"
181..3054
/note="unnamed protein product; dystrophin related protein match=807)
2 (MGD|MG1:107432, GB|U43520, evidence: BLASTN, 99%, putative)
/codon_start=1
/protein_id="BAC38217.1"
/db_xref="GI:26349155"
/translation="MQPLVMQGPYTLPRCHEWHAADRFHHSSSLRNTCPQPQVRAAV TIPAPPWDGDPCLSPKLLNGTVGATGPLEPSAMNLCWNEIKKSHNLRARLEAFSD LSGKLQPLREIIDLWSQDELSAQLPLQGDVALVQOEKETHAAFMEEVSKSGPYIS SVLESAQAFILSQHPFEELESQSESKDTPRQIQLNLSRFVWKQATVASSELWEKLTAR CVDQHRHIEHLEHLEIQGAMEELESLTLQAEGVRATWEPIDGLFIDSLPEHIQAIK LFKEEFPVKDGVKLVNDLAHQLAISDVHLSMENSRALEQINIRWKQLQVSAERLKQ LQDAHRDFGSGHQHFLSTSVQVPWERAFSPNKVPYYINHQAQTCWDHPKMTLYOTL ADLNNIKFSAYRTAMKLRVQKALRLDLVLTLEIFNEHDLQASEHVMVDVVEVIHC LTALYERLEERGILVNPLCVDMSLNWLNVFDSGRGKMRALSFKTGIACLCGTEV KEKLQYLFQVANSQSCDQRLHLLGALLHEAIQVPRQLGEVAAFSGSNVPSVRSCTFRF STGKPIVEASQFLEWNLPEQSMVLAHLRVITIAEQVHQTCKSICRQCPIKGPYR SLKFQNVDCIQCTFLTRASKGNKLYHPIMEYVYPTTSSENVDFATTLKNKFRKQY FSKHPORGVLVQSVLESDCSETPASSPMLPHADTHSRIEHPASRLAEMESONCSFFN DLSLPDSDIDEDQYLLRHSSPITDREPAFGQAPCSMATESKGELEKILAHLEDENRI LQGLRLKWOHEEAAPTLLVEGSAEATPDHRNEELAEARILRQHKSRLETRMOIL EDHNNKQLESQRLRELLELLOPPSESDGNGSAGSSLASPPROSEGSHPREKGGTTDTE VADDVGSKSQDVSLCLEDIMEKLRHAFPSVRSDDVTANTILLAS"

CDS

3731..3736
/note="putative"
3753
/note="putative"
/note="putative"
927 a 1021 c 938 g 867 t

polyA_signal

polyA_site

| | | | | |
|-----------------------|---|--------------------------------|---------------------------------------|---------|
| BASE COUNT | 927 a | 1021 c | 938 g | 867 t |
| ORIGIN | | | | |
| Query Match | 15.5%; Score 309.8; DB 11; Length 3753; | | | |
| Best Local Similarity | 59.1%; Pred. No. 4.2e-52; | | | |
| Matches | 530; Conservative | 0; Mismatches | 367; Indels | 0; Gaps |
| 0; | | | | |
| QY | 1104 | CATGAACCTCAAGTGGAGTGAACCTTCGGA | AAAGTCTCTCAACATTAGTCCCATTTGGA | 1163 |
| Db | 414 | CATGAATCTGTGTTGGAATGAATAA | AAAGAGTCTCACAACCTCCGCGCTCGCCTAGA | 473 |
| QY | 1164 | AGCCAGTTCTGACCAGTGGAGCGTCTGC | ACCCTTCTCTGCAGGAACCTCTGGTGTGGCT | 1223 |
| Db | 474 | GGCCTTCTCAGACCTCAGTGGAAAAC | TTTCAGCTCCCGTCCGAGAGATTATTGACTGGCT | 533 |
| QY | 1224 | ACAGTGAAAGATGATGATTAAGCCGCG | CAGGCACCTATTGGAGGCGACTTCCAGCAGT | 1283 |
| Db | 534 | CAGCAAAAGGATGAGGAGTTGTCTC | AGCTCAGCTCCCTTGCAGAGGGATGTGGCCCTGGT | 593 |
| QY | 1284 | TCAGAGCAGAACGATGTACATAGGCG | CTTCAAGAGGGAATTGAAACCTAAGAACCTGT | 1343 |
| Db | 594 | ACAAACAGGAGAAGGAGACACATGC | AGCCTTTCATGGAAGAAGTCAAGTCTAAGGGCCCTTA | 653 |
| QY | 1344 | AATCATGAGTACTCTTGTGAGACTGT | TACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACT | 1403 |
| Db | 654 | CATCTCCTCTGTGCTTGAATCCGCC | CAGGCTTTTCTGTCTCCAGCATCCATTGGAAGAATT | 713 |

| | | | |
|----|------|---|------|
| QY | 1404 | AGAGAAACTCTACAGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAAATGTCAC | 1463 |
| Db | 714 | AGAGGAGTCTCAATCTGAGAGTAAAGATACTCCCCAGACAGCGGATTCAAACCTTAG | 773 |
| QY | 1464 | TCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCA | 1523 |
| Db | 774 | TCGCTTTGTATGGAAGCAGGCAACAGTGGCCAGTGAGCTGTGGGAGAAGCTGACAGCCCG | 833 |
| QY | 1524 | CTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAAGACTCCAGGAACCTCAAGA | 1583 |
| Db | 834 | CTGTGTAGATCAGCATCGCCACATTGAGCATACTCTGGAACATCTATTGGAGATCCAAGG | 893 |
| QY | 1584 | GGCCACGGATGAGTGGACCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCA | 1643 |
| Db | 894 | GGCAATGGAGGAATTGAGCAGTACTTTGACCCAAGCAGAGGGAGTCCGAGCCACGTGGGA | 953 |
| QY | 1644 | GCCCGTGGCGATCTCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACT | 1703 |
| Db | 954 | GCCCATAGGAGATCTCTTTATTGATTCCCTCCAGAGCATATCCAAGCCATCAAGCTATT | 1013 |
| QY | 1704 | TCGAGGAGAAAATTGCGCCTCTGAAAGAGAACGTGAGGCCACGTCAATGACCTTGCTGCCA | 1763 |
| Db | 1014 | CAAAGAAGAATTCTCTCTGTGAAAGATGGGGTGAAGTTAGTGAATGATCTGGCCCCACCA | 1073 |
| QY | 1764 | GCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAAACCCTCAGCACTCTGGAAGACCTGAA | 1823 |
| Db | 1074 | GCTTGCCATCTCTGATGTGCACCTTGTCATATGGAGAATTCAGGGCTCTTGGAAACAGATCAA | 1133 |
| QY | 1824 | CACCAGATGGAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAGGAGCTGCATGAAGC | 1883 |
| Db | 1134 | CATCCGGTGAAACAGCTCCAGGTGTGAGTGGCTGAGAGGCTTAAGCAACTCCAGGATGC | 1193 |
| QY | 1884 | CCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTGTCCAGGGTCCCTG | 1943 |
| Db | 1194 | CCACCGGACTTTGGCGCTGGGTACAGCACTTCTCTCCACTTCTGTCCAAGTTCCCTG | 1253 |
| QY | 1944 | GGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAAC | 2000 |
| Db | 1254 | GGAAAGAGCATTTTCTCCCAATAAAGTCCCTACTACATCAACCACGAGGCTCAGAC | 1310 |

RESULT 13

| | | | | | | |
|------------|---|----------|---------|------|--------|-----------------|
| BC036095 | LOCUS | BC036095 | 3051 bp | mRNA | linear | HTC 04-MAR-2003 |
| DEFINITION | Homo sapiens, clone IMAGE:5300319, mRNA. | | | | | |
| ACCESSION | BC036095 | | | | | |
| VERSION | BC036095.1 GI:23271908 | | | | | |
| KEYWORDS | HTC. | | | | | |
| SOURCE | Homo sapiens (human) | | | | | |
| ORGANISM | Homo sapiens | | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | | |
| | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | | |
| | 1 (bases 1 to 3051) | | | | | |
| | Strausberg,R. | | | | | |
| AUTHORS | Direct Submission | | | | | |
| | Submitted (31-JUL-2002) National Institutes of Health, Mammalian | | | | | |
| | Gene Collection (MGC), Cancer Genomics Office, National Cancer | | | | | |
| | Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, | | | | | |
| REMARK | USA | | | | | |
| | NIH-MGC Project URL: http://mgc.nci.nih.gov | | | | | |
| | Contact: MGC help desk | | | | | |
| | Email: cgapbs-remail.nih.gov | | | | | |
| COMMENT | Tissue Procurement: Miklos Palkovits, M.D., Ph.D. | | | | | |
| | cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki | | | | | |
| | Toshiyuki and Piero Carninci (RIKEN) | | | | | |
| | cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) | | | | | |
| | DNA Sequencing by: Sequencing Group at the Stanford Human Genome | | | | | |
| | Center, Stanford University School of Medicine, Stanford, CA 94305 | | | | | |
| | Web site: http://www-shgc.stanford.edu | | | | | |
| | Contact: (Dickson, Mark) mcd@paxil.stanford.edu | | | | | |
| | Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, | | | | | |
| | R. M. | | | | | |

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 48 Row: 0 Column: 2
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4503392
This clone has the following problem: retained intron.

FEATURES

source
1. .3051
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5300319"
/tissue_type="Brain, hypothalamus"
/clone_lib="NIH_MGC_96"
/lab_host="DH10B"
/note="Vector: pBluescript"
BASE COUNT 848 a 724 c 725 g 754 t
ORIGIN

Query Match 13.3%; Score 265.4; DB 11; Length 3051;
Best Local Similarity 57.8%; Pred. No. 3.8e-43;
Matches 473; Conservative 0; Mismatches 346; Indels 0; Gaps 0;
2Y 1105 ATGAACCTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAA 1164
2b 553 ATGAATCTGTGTTGGAATGAATAAAAAAGAGTCTCACAACTCCGGCGTCGCCTAGAG 612
2Y 1165 GCCAGTTCTGACCAGTGGAAAGCCTCTGCACCTTTCTTCAGGAACCTTCGTGTGGCTA 1224
2b 613 GCCTTCTCAGACCACACAGTGGAAAAGCTTCAGTCCCTCTTCAAGAGATTATTGACTGGCTC 672
2Y 1225 CAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTTANTGGAGGCGACTTTCAGCAGTT 1284
2b 673 AGCCAAAAGGATGAGGAGTTGTGAGCTCAGCTGCCCTCTACAGGGGGATGTGGSCCTGGTG 732
2Y 1285 CAGAAAGACAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACATAAAGAACCTGTA 1344
2b 733 CAACAGGAGAAGGAGACACATCGGGCCTTTATGGAAGAAGTCAAGTCTCGGGGCCCTAC 792
2Y 1345 ATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTA 1404
2b 793 ATCTATTCTGTGTGGAGTCAGTCAAGTCAAGCCTTCCTGTCCAGCACCCATTGAGGAGTTA 852
2Y 1405 GAGAAACTCTACCAGGAGCCCAAGAGAGCTGCCCTCTGAGGAGAGAGCCCAAGATGTCACT 1464
2b 853 GAGGAGCTCATTTCTGAGAGCAAGAATACCTCCCCGAAACAGCGGATCCAGAATCTCAGC 912
2Y 1465 CGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCAC 1524
2b 913 CGCTTTGTATGGAAGCAGGCGGACGGTGGCCAGTGAAGTGTGGGAGAAGTTGACAGCCCGC 972
2Y 1525 TCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTCAAGAG 1584
2b 973 TGTGTGACCAAGCACCGTCAATTGAGCGGACTCTGGAGCAGCTCTTGGAGATTCAAGGG 1032
2Y 1585 GCCACGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCTCTGGCAG 1644
2b 1033 GCAATGGAGGAACATAAGCACTACTCTGAGCCCAAGCTGAGGGAGTCCGAGCCACTTGGGAG 1092
2Y 1645 CCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTT 1704
2b 1093 CCCATTGGGGATCTCTTCATTGATTCACTCCAGAGCACATCCAGGCTATTAAAGCTGTTT 1152
2Y 1705 CGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAG 1764
2b 1153 AAAGAGAAATTCTCCCCCATGAAGATGGAGTAAAGTTGGTGAATGATCTGGGCCACCAA 1212
2Y 1765 CTTACCACTTTGGGCATTCAGCTCTCAACCGTATAACCTCAGCACTCTGGAAGACCTGAAC 1824
2b 1213. CTTGCCATTTCTGATGTGCACTTGTCAATGGAGAATTCCAGGCCCTTGGAAACAGATCAAC 1272
2Y 1825 ACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGTCATGAAGCC 1884

Db 1273 GTCCGATGGAACAACTACAGGCGTCAGTTAGTGAGAGGCTTAAGCAGCTCCAGGATGCC 1332
QY 1885 CACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCC 1923
Db 1333 CACCGGACTTTGGGCTGGGTACAGCACTTTCTCTCC 1371

RESULT 14
BC011062
LOCUS BC011062 2334 bp mRNA linear HTC 04-MAR-2003
DEFINITION Mus musculus, similar to utrophin, clone IMAGE:3979320, mRNA.
ACCESSION BC011062
VERSION BC011062.1 GI:15029695
KEYWORDS HTC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2334)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (25-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunsaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Lounseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 23 Row: j Column: 4
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein
This clone has the following problem: retained intron.

FEATURES

source
1. .2334
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3979320"
/tissue_type="Mammary tumor. Metallothionien-TGF alpha
model. 10 month old virgin mouse. Taken by biopsy."
/clone_lib="NCI CGAP_Mam1"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
BASE COUNT 698 a 512 c 638 g 486 t
ORIGIN

Query Match 13.1%; Score 261.2; DB 11; Length 2334;
Best Local Similarity 61.7%; Pred. No. 2.6e-42;
Matches 416; Conservative 0; Mismatches 258; Indels 0; Gaps 0;

QY 4 AGTTCAATGATGGAGAGTGAAGTAAACCTGGACCGCTTATCAACAGCTTTAGAGAAGTA 63
Db 1042 AGCACCGTCACTGAAGTGGACATGGATTTGGACAGCTACCAGATAGCGCTAGAGGAAGTG 1101
QY 64 TTATCGTGGCTTCTTTCTGTGAGGACACATTTGCAAGCACCAAGGAGAGATTTCTAATGAT 123

Db 1102 CTGACGTGGCTGTCTCCGGCAGGACACGTTCCAGGAGCAAGATGACATTTCTGTATGAT 1161

Qy 124 GTGGAAGTGGTGAAGACCAAGTTTCATCTCATGAGGGGTACATGATGGATTGACAGCC 183

Db 1162 GTCGAAGAAGTCAAGAGCAGTTTGCTACCCATGAAACTTTTATGATGGAGCTGACAGCA 1221

Qy 184 CATCAGGCCGGGTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAA 243

Db 1222 CACCAGAGCAGCTGGGGAGCGTCTCGCAGGCTGGCAACCAAGCTGATGACACAAGGGACT 1281

Qy 244 TTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATCAAGATGG 303

Db 1282 CTGTCAGAGGAGGAGGAGTTTGAGATCCAGGAACAGATGACCTTGCTGAATGCAAGGTGG 1341

Qy 304 GAATGCCTCAGGGTAGCTAGCATGGAAAGCAATTTACATAGAGTTTAAATGGAT 363

Db 1342 GAGGCGCTCCGGTGGAGAGGATGGAGAGGCGAGTCCCGGCTGCACGACGCTCTGTATGGAG 1401

Qy 364 CTCCACAATCAGAAACTGAACAGATTGAATGACTGCTGCTAACAACAAACAGAGAAACA 423

Db 1402 CTGCAGAAGAAACAGCTGCAGCAGCTCTCAAGCTGGCTGGCCCTCACAGAGAGCGCATT 1461

Qy 424 AGGAATAATGGAGGAAGAGCCCTCTTGGACCTGTATCTGAAGACCTAAACCGCAAGTACAA 483

Db 1462 CAGAAGATGGAGAGAGCCCTCCCGTGGTGATGACCTCGCCCTCCCTGCAGAGCTGCTTCAA 1521

Qy 484 CAACATAAGGTGCTTCAAGAGATCTAGAACAAAGACAAAGTCAGGCTCAATTTCTCTCACT 543

Db 1522 GAACATAAAAGTTTGCAAAATGACCTTGAAGCTGAACAGGTGAAGGTAATTCCTTAACT 1581

Qy 544 CACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGGAAGAA 603

Db 1582 CACATGGTGGTGATTGTGGATGAAACACAGTGGGAGAGTGCCACAGCTCTTCTGGAAGAT 1641

Qy 604 CAACTTAAGGTATTGGGAGATCGATGGGCAACATCTCTAGATGGACAGAACCGCTGG 663

Db 1642 CAGTTACAGAAACTGGTGAGCGCTGGACAGCTGTATGCCGCTGGACTGAAGAACGTTGG 1701

Qy 664 GTTCTTTTACAAGA 677

Db 1702 AACAGGTTGCAAGA 1715

RESULT 15

AK034383

LOCUS

DEFINITION

AK034383 1541 bp mRNA linear HTC 05-DEC-2002

Mus musculus adult male diencephalon cDNA, RIKEN full-length enriched library, clone:9330185F14 product:dystrophin related protein 2, full insert sequence.

ACCESSION

AK034383

VERSION

AK034383.1 GI:26329906

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)

TITLE

High-efficiency full-length cDNA cloning

JOURNAL

Meth. Enzymol. 303, 19-44 (1999)

MEDLINE

99279253

PUBMED

10349636

REFERENCE

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

TITLE

Prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL

Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE

20499374

PUBMED

11042159

REFERENCE

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleschmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection

Nature 409 (6821), 685-690 (2001)

21085660

11217851

5

The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 1541)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL:http://genome.gsc.riken.go.jp/

URL:http://fantom.gsc.riken.go.jp/.

Location/Qualifiers

1. 1541

/organism="Mus musculus"

/mol_type="mRNA"

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

2M nucleic - nucleic search, using sw model

run on: February 1, 2004, 23:51:42 ; Search time 655.123 Seconds
(without alignments)
11131.886 Million cell updates/sec

Title: US-09-845-416-6_COPY_1000_3000

Perfect score: 2001

Sequence: 1 ggcagttcattgatggagag.....tcaaccacgagactcaaca 2001

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2434939 seqs, 1822278265 residues

Total number of hits satisfying chosen parameters: 4869878

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|--------|---------------|--------|-------|-------------------------------------|
| 1 | 2001 | 100.0 | 3999 | 13 | US-09-845-416-6 Sequence 6, Appli |
| 2 | 2001 | 100.0 | 4966 | 13 | US-09-845-416-28 Sequence 28, Appl |
| 3 | 2001 | 100.0 | 4990 | 13 | US-09-845-416-34 Sequence 34, Appl |
| 4 | 1709 | 85.4 | 3858 | 13 | US-09-845-416-9 Sequence 9, Appli |
| 5 | 1709 | 85.4 | 4825 | 13 | US-09-845-416-29 Sequence 29, Appl |
| 6 | 1709 | 85.4 | 4848 | 13 | US-09-845-416-35 Sequence 35, Appl |
| 7 | 1709 | 85.4 | 5060 | 13 | US-09-845-416-36 Sequence 36, Appl |
| 8 | 1662.8 | 83.1 | 4182 | 13 | US-09-845-416-2 Sequence 2, Appli |
| 9 | 1662.8 | 83.1 | 5149 | 13 | US-09-845-416-27 Sequence 27, Appl |
| 10 | 1283 | 64.1 | 5462 | 13 | US-10-149-736-41 Sequence 41, Appl |
| 11 | 1182.6 | 59.1 | 8689 | 13 | US-10-149-736-42 Sequence 42, Appl |
| 12 | 1182.6 | 59.1 | 11058 | 13 | US-09-845-416-1 Sequence 1, Appli |
| 13 | 1182.6 | 59.1 | 11443 | 13 | US-10-149-736-44 Sequence 44, Appl |
| 14 | 1182.6 | 59.1 | 12057 | 13 | US-10-149-736-47 Sequence 47, Appl |
| 15 | 1182.6 | 59.1 | 13957 | 10 | US-09-782-378A-22 Sequence 22, Appl |

| | | | | | |
|----|--------|------|-------|----|---------------------------------------|
| 16 | 1182.6 | 59.1 | 13957 | 10 | US-09-880-107-2284 Sequence 2284, Ap |
| 17 | 1182.6 | 59.1 | 13957 | 13 | US-10-149-736-1 Sequence 1, Appli |
| 18 | 1182.6 | 59.1 | 14082 | 13 | US-10-341-434-108 Sequence 108, App |
| 19 | 1182 | 59.1 | 2169 | 13 | US-09-845-416-4 Sequence 4, Appli |
| 20 | 1182 | 59.1 | 3531 | 13 | US-09-845-416-10 Sequence 10, Appl |
| 21 | 1182 | 59.1 | 4498 | 13 | US-09-845-416-30 Sequence 30, Appl |
| 22 | 1180 | 59.0 | 5339 | 13 | US-10-149-736-40 Sequence 40, Appl |
| 23 | 1013 | 50.6 | 3510 | 13 | US-09-845-416-12 Sequence 12, Appl |
| 24 | 1013 | 50.6 | 4476 | 13 | US-09-845-416-31 Sequence 31, Appl |
| 25 | 1004 | 50.2 | 13815 | 13 | US-10-149-736-2 Sequence 2, Appli |
| 26 | 835 | 41.7 | 1821 | 13 | US-09-845-416-13 Sequence 13, Appl |
| 27 | 727 | 36.3 | 5417 | 13 | US-10-149-736-39 Sequence 39, Appl |
| 28 | 678.4 | 33.9 | 4414 | 13 | US-09-845-416-32 Sequence 32, Appl |
| 29 | 678 | 33.9 | 1991 | 13 | US-09-845-416-3 Sequence 3, Appli |
| 30 | 677 | 33.8 | 1667 | 13 | US-09-845-416-7 Sequence 7, Appli |
| 31 | 666.4 | 33.3 | 3446 | 13 | US-09-845-416-14 Sequence 14, Appl |
| 32 | 448 | 22.4 | 1434 | 13 | US-09-845-416-15 Sequence 15, Appl |
| 33 | 408.2 | 20.4 | 11096 | 13 | US-10-149-736-4 Sequence 4, Appli |
| 34 | 397 | 19.8 | 10302 | 10 | US-09-782-378A-23 Sequence 23, Appl |
| 35 | 397 | 19.8 | 10302 | 13 | US-10-149-736-3 Sequence 3, Appli |
| 36 | 387 | 19.3 | 387 | 13 | US-10-149-736-32 Sequence 32, Appl |
| 37 | 385 | 19.2 | 16531 | 13 | US-10-101-510-667 Sequence 667, App |
| 38 | 350 | 17.5 | 1340 | 13 | US-09-845-416-11 Sequence 11, Appl |
| 39 | 348 | 17.4 | 348 | 13 | US-10-149-736-31 Sequence 31, Appl |
| 40 | 330 | 16.5 | 333 | 13 | US-10-149-736-9 Sequence 9, Appli |
| 41 | 327 | 16.3 | 327 | 13 | US-10-149-736-8 Sequence 8, Appli |
| 42 | 324 | 16.2 | 324 | 13 | US-10-149-736-33 Sequence 33, Appl |
| 43 | 160.2 | 8.0 | 256 | 9 | US-09-864-761-21956 Sequence 21956, A |
| 44 | 153.4 | 7.7 | 467 | 9 | US-09-864-761-11083 Sequence 11083, A |
| 45 | 151 | 7.5 | 151 | 9 | US-09-864-761-27715 Sequence 27715, A |

ALIGNMENTS

RESULT 1

US-09-845-416-6
; Sequence 6, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 6
; LENGTH: 3999
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-6

| | | | | |
|-----------------------|--------------|---|------------|--------------|
| Query Match | 100.0%; | Score 2001; | DB 13; | Length 3999; |
| Best Local Similarity | 100.0%; | Pred. No. 0; | | |
| Matches 2001; | Conservative | 0; | Mismatches | 0; |
| | | | Indels | 0; |
| | | | Gaps | 0; |
| QY | 1 | GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA | 60 | |
| | | | | |
| Db | 1000 | GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA | 1059 | |
| | | | | |
| QY | 61 | GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGCAAGGAGAGATTCTTAAT | 120 | |
| | | | | |
| Db | 1060 | GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGCAAGGAGAGATTCTTAAT | 1119 | |
| | | | | |
| QY | 121 | GATGTGGAAGTGGTGAAGACACAGTTTTCATCTACTATGAGGGGTACATGATGATTGACA | 180 | |
| | | | | |
| Db | 1120 | GATGTGGAAGTGGTGAAGACACAGTTTTCATCTACTATGAGGGGTACATGATGATTGACA | 1179 | |
| | | | | |
| QY | 181 | GCCCATCAGGGCCGGTGGTAAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGA | 240 | |

Db 1180 GCCCATCAGGGCCGGTTGGTAAATATTCTACAATTTGGGAAGTAAGCTGATTGGAAACAGGA 1239
QY 241 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCTTAATTTCAAGA 300
Db 1240 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCTTAATTTCAAGA 1299
QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATG 360
Db 1300 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATG 1359
QY 361 GATCTCCAGAAATCAGAAACTGAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGA 420
Db 1360 GATCTCCAGAAATCAGAAACTGAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGA 1419
QY 421 ACAAGGAAAAATGGAGGAAGAGCCTCTTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 480
Db 1420 ACAAGGAAAAATGGAGGAAGAGCCTCTTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 1479
QY 481 CAACAACATAAAGGTGCTTCAAGAAGATCTAGAACAAGAACTAGGTCAGGGTCAATTTCTCTC 540
Db 1480 CAACAACATAAAGGTGCTTCAAGAAGATCTAGAACAAGAACTAGGTCAGGGTCAATTTCTCTC 1539
QY 541 ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 600
Db 1540 ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 1599
QY 601 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCCGC 660
Db 1600 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCCGC 1659
QY 661 TGGGTTCTTTTACAAGACCAGCCTGACCTAGCTCCTGGACTGACCACCTATTGGAGCCTCT 720
Db 1660 TGGGTTCTTTTACAAGACCAGCCTGACCTAGCTCCTGGACTGACCACCTATTGGAGCCTCT 1719
QY 721 CCTACTCAGACTGTTACTCTGCTGACACAACTGTGTTACTAAGGAAACTGSCATCTCC 780
Db 1720 CCTACTCAGACTGTTACTCTGCTGACACAACTGTGTTACTAAGGAAACTGSCATCTCC 1779
QY 781 AAACHTAGAAATGCCATCTTCCCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 840
Db 1780 AAACHTAGAAATGCCATCTTCCCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 1839
QY 841 TTCCCCCTGGACCTGGAAAAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACACTGCCAAT 900
Db 1840 TTCCCCCTGGACCTGGAAAAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACACTGCCAAT 1899
QY 901 GTCTTACAGATGCTACCCGTAAGAAAGGCTCCTAGAGACTCCAAGGAGTAAGAGAG 960
Db 1900 GTCTTACAGATGCTACCCGTAAGAAAGGCTCCTAGAGACTCCAAGGAGTAAGAGAG 1959
QY 961 CTGATGAACAATGGCAAGACCTCCAAGGTGAAATTGAAAGCTCACACAGATGTTTATCAC 1020
Db 1960 CTGATGAACAATGGCAAGACCTCCAAGGTGAAATTGAAAGCTCACACAGATGTTTATCAC 2019
QY 1021 AACCTGGATGAAAAACAGCCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTC 1080
Db 2020 AACCTGGATGAAAAACAGCCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTC 2079
QY 1081 CTGTTACAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTCGGAAAAAGTCT 1140
Db 2080 CTGTTACAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTCGGAAAAAGTCT 2139
QY 1141 CTCAACATTAGGTCCCATTTTGGAAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCT 1200
Db 2140 CTCAACATTAGGTCCCATTTTGGAAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCT 2199
QY 1201 CTGCAGGAACCTTCTGGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT 1260
Db 2200 CTGCAGGAACCTTCTGGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT 2259
QY 1261 ATTGGAGGCGACTTTCAGCAGTTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGG 1320

Db 2260 ATTGAGGGCGACTTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGG 2319
QY 1321 GAATTGAAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 1380
Db 2320 GAATTGAAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 2379
QY 1381 ACAGAGCAGCCTTTTGGAAAGACTAGAGAAAACTTACAGGAGCCCAGAGAGTGCCTCTCT 1440
Db 2380 ACAGAGCAGCCTTTTGGAAAGACTAGAGAAAACTTACAGGAGCCCAGAGAGTGCCTCTCT 2439
QY 1441 GAGGAGAGAGCCCCAGAAATGTCTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 1500
Db 2440 GAGGAGAGAGCCCCAGAAATGTCTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 2499
QY 1501 GAGTGGGAAAAATTTGAACTCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 1560
Db 2500 GAGTGGGAAAAATTTGAACTCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 2559
QY 1561 GAAAGACTCCAGGAACCTTCAAGAGGGCCACGGATGAGCTGGACCTCAAGCTGCGGCCAAGCT 1620
Db 2560 GAAAGACTCCAGGAACCTTCAAGAGGGCCACGGATGAGCTGGACCTCAAGCTGCGGCCAAGCT 2619
QY 1621 GAGTGTATCAAGGGATCCTTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAAGAT 1680
Db 2620 GAGTGTATCAAGGGATCCTTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAAGAT 2679
QY 1681 CACTTCGAGAAAGTCAAGSCACTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTTGAGC 1740
Db 2680 CACTTCGAGAAAGTCAAGSCACTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTTGAGC 2739
QY 1741 CAGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAAC 1800
Db 2740 CAGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAAC 2799
QY 1801 CTGAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGAC 1860
Db 2800 CTGAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGAC 2859
QY 1861 CGAGTCAGGCAGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTT 1920
Db 2860 CGAGTCAGGCAGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTT 2919
QY 1921 TCCAGTCTGTCCAGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTAT 1980
Db 2920 TCCAGTCTGTCCAGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTAT 2979
QY 1981 ATCAACCACGAGACTCAAAACA 2001
Db 2980 ATCAACCACGAGACTCAAAACA 3000

RESULT 2
US-09-845-416-28
; Sequence 28, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 28
; LENGTH: 4966
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-28

Query Match 100.0%; Score 2001; DB 13; Length 4966;

| | | | | | | | |
|--|------|---|------|---|--|--|--|
| Best Local Similarity 100.0%; Pred. No. 0; | | | | Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | |
| QY | 1 | GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAA | 60 | | | | |
| Db | 1757 | GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAA | 1816 | | | | |
| QY | 61 | GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCCTAAT | 120 | | | | |
| Db | 1817 | GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCCTAAT | 1876 | | | | |
| QY | 121 | GATGTGGAAGTGGTGAAAGACCAGTTTCATACTCATGAGGGGTACATGATGGATTGACA | 180 | | | | |
| Db | 1877 | GATGTGGAAGTGGTGAAAGACCAGTTTCATACTCATGAGGGGTACATGATGGATTGACA | 1936 | | | | |
| QY | 181 | GCCCATCAGGCGCGGTTGGTAATATCTACAAATTGGGAAGTAAAGCTGATTGGAAACAGGA | 240 | | | | |
| Db | 1937 | GCCCATCAGGCGCGGTTGGTAATATCTACAAATTGGGAAGTAAAGCTGATTGGAAACAGGA | 1996 | | | | |
| QY | 241 | AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA | 300 | | | | |
| Db | 1997 | AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA | 2056 | | | | |
| QY | 301 | TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG | 360 | | | | |
| Db | 2057 | TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG | 2116 | | | | |
| QY | 361 | GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAGA | 420 | | | | |
| Db | 2117 | GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAGA | 2176 | | | | |
| QY | 421 | ACAAGGAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTA | 480 | | | | |
| Db | 2177 | ACAAGGAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTA | 2236 | | | | |
| QY | 481 | CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTCCTC | 540 | | | | |
| Db | 2237 | CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTCCTC | 2296 | | | | |
| QY | 541 | ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAAGCAACTGCTGCTTTGGAA | 600 | | | | |
| Db | 2297 | ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAAGCAACTGCTGCTTTGGAA | 2356 | | | | |
| QY | 601 | GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGC | 660 | | | | |
| Db | 2357 | GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGC | 2416 | | | | |
| QY | 661 | TGGGTTCTTTTACAAGACCAGCCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTCT | 720 | | | | |
| Db | 2417 | TGGGTTCTTTTACAAGACCAGCCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTCT | 2476 | | | | |
| QY | 721 | CCTACTCAGACTGTTACTCTGGTGACAAACCTGTGGTTACTAAGGAAACTGCCATCTCC | 780 | | | | |
| Db | 2477 | CCTACTCAGACTGTTACTCTGGTGACAAACCTGTGGTTACTAAGGAAACTGCCATCTCC | 2536 | | | | |
| QY | 781 | AAACTAGAAATGCCATCTTCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG | 840 | | | | |
| Db | 2537 | AAACTAGAAATGCCATCTTCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG | 2596 | | | | |
| QY | 841 | TTCCCCCTGGACCTGGAAAAGTTTCTTGCCCTGGCTTACAGAAGCTGAAACAACTGCCAAT | 900 | | | | |
| Db | 2597 | TTCCCCCTGGACCTGGAAAAGTTTCTTGCCCTGGCTTACAGAAGCTGAAACAACTGCCAAT | 2656 | | | | |
| QY | 901 | GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAAGCTCCAAGGGAGTAAAGAG | 960 | | | | |
| Db | 2657 | GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAAGCTCCAAGGGAGTAAAGAG | 2716 | | | | |
| QY | 961 | CTGATGAAACAAATGGCAAGACCCTCCNAGGTGAAATTGAAGCTCACACAGATGTTTATCAC | 1020 | | | | |
| Db | 2717 | CTGATGAAACAAATGGCAAGACCCTCCNAGGTGAAATTGAAGCTCACACAGATGTTTATCAC | 2776 | | | | |
| QY | 1021 | AACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAGGTTCCGATGATGCAGTC | 1080 | | | | |

| | | | | | | | |
|----|------|--|------|--|--|--|--|
| Db | 2777 | AACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTC | 2836 | | | | |
| QY | 1081 | CTGTTTCAAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACTTCGGAAAAAGTCT | 1140 | | | | |
| Db | 2837 | CTGTTTCAAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACTTCGGAAAAAGTCT | 2896 | | | | |
| QY | 1141 | CTCAACATTAGGTCCCATTTTGGAGCCAGTTTCTGACCAGTGGAAAGCGTCTGCACCTTTCT | 1200 | | | | |
| Db | 2897 | CTCAACATTAGGTCCCATTTTGGAGCCAGTTTCTGACCAGTGGAAAGCGTCTGCACCTTTCT | 2956 | | | | |
| QY | 1201 | CTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAAGATGATGAATTAAGCCGGCAGGCACCT | 1260 | | | | |
| Db | 2957 | CTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAAGATGATGAATTAAGCCGGCAGGCACCT | 3016 | | | | |
| QY | 1261 | ATTGGAGGCGACTTTCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGG | 1320 | | | | |
| Db | 3017 | ATTGGAGGCGACTTTCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGG | 3076 | | | | |
| QY | 1321 | GAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTG | 1380 | | | | |
| Db | 3077 | GAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTG | 3136 | | | | |
| QY | 1381 | ACAGAGCAGCCTTTTGGAAAGGACTAGAGAAAACTCTACCAGGAGCCCAGAGAGTGCCTCCT | 1440 | | | | |
| Db | 3137 | ACAGAGCAGCCTTTTGGAAAGGACTAGAGAAAACTCTACCAGGAGCCCAGAGAGTGCCTCCT | 3196 | | | | |
| QY | 1441 | GAGGAGAGAGCCCAAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT | 1500 | | | | |
| Db | 3197 | GAGGAGAGAGCCCAAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT | 3256 | | | | |
| QY | 1501 | GAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT | 1560 | | | | |
| Db | 3257 | GAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT | 3316 | | | | |
| QY | 1561 | GAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGAACCTCAAGCTGCGCCCAAGCT | 1620 | | | | |
| Db | 3317 | GAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGAACCTCAAGCTGCGCCCAAGCT | 3376 | | | | |
| QY | 1621 | GAGGTGATCAAGGGATCCTGGAGAGCCCGTGGCGATCTCTCAATTGACTCTCTCCAAGAT | 1680 | | | | |
| Db | 3377 | GAGGTGATCAAGGGATCCTGGAGAGCCCGTGGCGATCTCTCAATTGACTCTCTCCAAGAT | 3436 | | | | |
| QY | 1681 | CACCTCGAGAAAGTCAAGGCATTCGAGGAGAAATTGCGCTCTGAAAGAGAACGTGAGC | 1740 | | | | |
| Db | 3437 | CACCTCGAGAAAGTCAAGGCATTCGAGGAGAAATTGCGCTCTGAAAGAGAACGTGAGC | 3496 | | | | |
| QY | 1741 | CACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTGAGCTCTCACCGTATAAC | 1800 | | | | |
| Db | 3497 | CACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTGAGCTCTCACCGTATAAC | 3556 | | | | |
| QY | 1801 | CTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGAC | 1860 | | | | |
| Db | 3557 | CTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGAC | 3616 | | | | |
| QY | 1861 | CGAGTCAGGCGAGCTGCATGAAGCCACAGGGACTTTTGGTCCAGCATCTCAGCACCTTTCTT | 1920 | | | | |
| Db | 3617 | CGAGTCAGGCGAGCTGCATGAAGCCACAGGGACTTTTGGTCCAGCATCTCAGCACCTTTCTT | 3676 | | | | |
| QY | 1921 | TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTCCCTACTAT | 1980 | | | | |
| Db | 3677 | TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTCCCTACTAT | 3736 | | | | |
| QY | 1981 | ATCAACCACGAGACTCAAACA | 2001 | | | | |
| Db | 3737 | ATCAACCACGAGACTCAAACA | 3757 | | | | |

RESULT 3
US-09-845-416-34
; Sequence 34, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO

; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 4990
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-34

Query Match 100.0%; Score 2001; DB 13; Length 4990;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA 60
1781 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA 1840
61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGTCAAGCACAAAGGAGAGATTTCTAAT 120
1841 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGTCAAGCACAAAGGAGAGATTTCTAAT 1900
121 GATGTGGAAGTGGTGAAGACACAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACA 180
1901 GATGTGGAAGTGGTGAAGACACAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACA 1960
181 GCCCATCAGGGCCGGTGGTAATATTCTACAATTGGGAAGTAAGCTGATTTGGAACAGGA 240
1961 GCCCATCAGGGCCGGTGGTAATATTCTACAATTGGGAAGTAAGCTGATTTGGAACAGGA 2020
241 AAATTATCAGAAGATGAAGAACTGAAGTACAAGACAGATGAATCTCTAAATTCAGA 300
2021 AAATTATCAGAAGATGAAGAACTGAAGTACAAGACAGATGAATCTCTAAATTCAGA 2080
301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATG 360
2081 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATG 2140
361 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGA 420
2141 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGA 2200
421 ACAAGGAAAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 480
2201 ACAAGGAAAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 2260
481 CAACAACATAAAGGTGCTTCAAGAAGATCTAGAACAGAAACAAGTCAGGTCAATTTCTCTC 540
2261 CAACAACATAAAGGTGCTTCAAGAAGATCTAGAACAGAAACAAGTCAGGTCAATTTCTCTC 2320
541 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 600
2321 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 2380
601 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGACCCG 660
2381 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGACCCG 2440
661 TGGGTTCTTTTACAAGACCCAGCTGACCTAGCTCCTGGACTGACCACCTATTGGAGCCTCT 720
2441 TGGGTTCTTTTACAAGACCCAGCTGACCTAGCTCCTGGACTGACCACCTATTGGAGCCTCT 2500
721 CCTACTCAGACTGTTACTCTGCTGACACAAACCTGTGGTTACTTAAGGAAACTGCCATCTCC 780
2501 CCTACTCAGACTGTTACTCTGCTGACACAAACCTGTGGTTACTTAAGGAAACTGCCATCTCC 2560
781 AAACTAGAAATGCCATCTTCTTGTGTTGGAGGTACCTACTCATAGATTACTGCAACAG 840

Db 2561 AAAC TAGAAATGCCATCTTCTTGATGTGGAGGTACCTACTCATAGATTACTGCAACAG 2620
QY 841 TTCCCTCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACTGCCAAT 900
Db 2621 TTCCCTCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACTGCCAAT 2680
QY 901 GTCCTACAGGATGTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAGAG 960
Db 2681 GTCCTACAGGATGTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAGAG 2740
QY 961 CTGATGAAACCAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCAC 1020
Db 2741 CTGATGAAACCAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCAC 2800
QY 1021 AACCTGGATGAAACAGCCAAAAATCCTGAGATCCCTGGAAGTTCGATGATGCAGTC 1080
Db 2801 AACCTGGATGAAACAGCCAAAAATCCTGAGATCCCTGGAAGTTCGATGATGCAGTC 2860
QY 1081 CTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCT 1140
Db 2861 CTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCT 2920
QY 1141 CTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCTCTGCACCTTTCT 1200
Db 2921 CTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCTCTGCACCTTTCT 2980
QY 1201 CTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAAGATGATGAATTAAGCCGCGAGGCACCT 1260
Db 2981 CTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAAGATGATGAATTAAGCCGCGAGGCACCT 3040
QY 1261 ATTGGAGGCGACTTCCAGCAGTTTCAGAGCAGAACGATGTACATAGGGCCTTCAAGAGG 1320
Db 3041 ATTGGAGGCGACTTCCAGCAGTTTCAGAGCAGAACGATGTACATAGGGCCTTCAAGAGG 3100
QY 1321 GAATTGAAAACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 1380
Db 3101 GAATTGAAAACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 3160
QY 1381 ACAGAGCAGCCTTTGGAAGGACTAGAGAACTCTACCAGGAGCCCGAGAGAGTGCCTCCT 1440
Db 3161 ACAGAGCAGCCTTTGGAAGGACTAGAGAACTCTACCAGGAGCCCGAGAGAGTGCCTCCT 3220
QY 1441 GAGGAGAGAGCCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 1500
Db 3221 GAGGAGAGAGCCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 3280
QY 1501 GAGTGGGAAAAATTGAACCTTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 1560
Db 3281 GAGTGGGAAAAATTGAACCTTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 3340
QY 1561 GAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCT 1620
Db 3341 GAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCT 3400
QY 1621 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGAT 1680
Db 3401 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGAT 3460
QY 1681 CACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTGAGC 1740
Db 3461 CACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTGAGC 3520
QY 1741 CACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAAC 1800
Db 3521 CACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAAC 3580
QY 1801 CTCAGCACTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGAGGTGGCGCTCGAGGAC 1860
Db 3581 CTCAGCACTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGAGGTGGCGCTCGAGGAC 3640
QY 1861 CGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTT 1920
Db 3641 CGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTT 3700

| | | | |
|----|------|--|------|
| QY | 1441 | GAGGAGAGAGCCCCAGAATGTCACTCGGCCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT | 1500 |
| | | | |
| Db | 3056 | GAGGAGAGAGCCCCAGAATGTCACTCGGCCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT | 3115 |
| | | | |
| QY | 1501 | GAGTGGGAAAAATTGAACCTTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTT | 1560 |
| | | | |
| Db | 3116 | GAGTGGGAAAAATTGAACCTTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTT | 3175 |
| | | | |
| QY | 1561 | GAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGGCCCAAGCT | 1620 |
| | | | |
| Db | 3176 | GAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGGCCCAAGCT | 3235 |
| | | | |
| QY | 1621 | GAGGTGATCAAGGATTCCTGGCAGCCCCTGGGCGAATCTCTCATTTGACTCTCTCCAAGAT | 1680 |
| | | | |
| Db | 3236 | GAGGTGATCAAGGATTCCTGGCAGCCCCTGGGCGAATCTCTCATTTGACTCTCTCCAAGAT | 3295 |
| | | | |
| QY | 1681 | CACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGAAAAATTGCGCCTCTGAAAGAGAACGTGAGC | 1740 |
| | | | |
| Db | 3296 | CACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGAAAAATTGCGCCTCTGAAAGAGAACGTGAGC | 3355 |
| | | | |
| QY | 1741 | CACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAAC | 1800 |
| | | | |
| Db | 3356 | CACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAAC | 3415 |
| | | | |
| QY | 1801 | CTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGAC | 1860 |
| | | | |
| Db | 3416 | CTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGAC | 3475 |
| | | | |
| QY | 1861 | CGAGTCAGGCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACATTTCCT | 1920 |
| | | | |
| Db | 3476 | CGAGTCAGGCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACATTTCCT | 3535 |
| | | | |
| QY | 1921 | TCCAGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT | 1980 |
| | | | |
| Db | 3536 | TCCAGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT | 3595 |
| | | | |
| QY | 1981 | ATCAACCCAGAGACTCAAACA | 2001 |
| | | | |
| Db | 3596 | ATCAACCCAGAGACTCAAACA | 3616 |
| | | | |

RESULT 6

US-09-845-416-35

Sequence 35. Application US/09845416

Publication No. US20030171312A1

; EDUCATION NO: 0020
; GENERAL INFORMATION:

APPLICANT: XIAO, XIAO

ALLEGENT: ARRO, ARRO
: TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE

FILE OF INVENTION: DATA CENTER
: TITLE OF INVENTION: THEREOF

FILE REFERENCE: DE1142

FILE REFERENCE: DETHIZ
: CURRENT APPLICATION NUMBER: US/0

CURRENT REFLECTION NUMBER: 03/03/95
CURRENT FILING DATE: 2001-04-30

CURRENT FILING DATE: 2001-04
PRIOR APPLICATION NUMBER: 60/

; PRIOR FILING DATE: 2000-
 ; PRIOR FILING DATE: 2000-

PREVIOUS FILING DATE: 2000-04-
: NUMBER OF SEC ID NOS: 36

; NUMBER OF SEQ
: SOFTWARE: Dat

```

; SOFTWARE: PAL
; CEO ID NO 35

```

; SEQ ID NO 3
1 ENCTH: 4; LENGTH: 4848
: TYPE: DNA; TYPE: DNA
OCCASION: HQ

```

Query Match      85.4%; Score 1709; DB 13; Length 4848;
Best Local Similarity 93.0%; Pred. No. 0;
Matches 1860; Conservative 0; Mismatches 0; Indels 141; Gaps 1;

```

| | | | | |
|----|------|--|-------------------|------|
| QY | 1 | GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAA | CAGCTTTAGAAGAA | 60 |
| | | | | |
| DB | 1780 | GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAA | CAGCTTTAGAAGAA | 1839 |
| | | | | |
| QY | 61 | GTATTATCGTGGCTTCCTTCTGCTGAGGACACATTGCAAGCA | CAGGAGAGATTCTTAAT | 120 |

```
QY 1201 CTGCAGGAACCTTCGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT 1260
Db 2839 CTGCAGGAACCTTCGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT 2898

QY 1261 ATTGAGGCGACTTCCAGCAGTTTCAGAACGAGAACGATGTACATAGGGCCCTTCAAGAGG 1320
Db 2899 ATTGAGGCGACTTCCAGCAGTTTCAGAACGAGAACGATGTACATAGGGCCCTTCAAGAGG 2958

QY 1321 GAATTGAAAACCTAAAGAACCTGTAAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTG 1380
Db 2959 GAATTGAAAACCTAAAGAACCTGTAAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTG 3018

QY 1381 ACAGAGCAGCCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCCGCAGAGAGCTGCCTCCT 1440
Db 3019 ACAGAGCAGCCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCCGCAGAGAGCTGCCTCCT 3078

QY 1441 GAGGAGAGAGCCCGAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 1500
Db 3079 GAGGAGAGAGCCCGAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 3138

QY 1501 GAGTGGGAAAAATTAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 1560
Db 3139 GAGTGGGAAAAATTAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 3198

QY 1561 GAAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCT 1620
Db 3199 GAAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCT 3258

QY 1621 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGAT 1680
Db 3259 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGAT 3318

QY 1681 CACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTGAGC 1740
Db 3319 CACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTGAGC 3378

QY 1741 CAGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGSCATTCAGCTCTCACCGTATAAC 1800
Db 3379 CAGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGSCATTCAGCTCTCACCGTATAAC 3438

QY 1801 CTCAGACTCTGGAAGACCTGAACACCATGGAAGCTTCTGAGGTGGCCGTCGAGGAC 1860
Db 3439 CTCAGACTCTGGAAGACCTGAACACCATGGAAGCTTCTGAGGTGGCCGTCGAGGAC 3498

QY 1861 CGAGTCAGGAGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTT 1920
Db 3499 CGAGTCAGGAGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTT 3558

QY 1921 TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT 1980
Db 3559 TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT 3618

QY 1981 ATCAACCACGAGACTCAAACA 2001
Db 3619 ATCAACCACGAGACTCAAACA 3639
```

RESULT 7

```
US-09-845-416-36
; Sequence 36, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 36
; LENGTH: 5060
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-36

Query Match      85.4%; Score 1709; DB 13; Length 5060;
Best Local Similarity 93.0%; Pred.No. 0;
Matches 1860; Conservative 0; Mismatches 0; Indels 141; Gaps 1;

QY 1 GGCAGTTCAATTGATGGAGAGTGAAGTAACCTGGACCGTTTATCAACACAGCTTTAGAAAGAA 60
Db 1992 GGCAGTTCAATTGATGGAGAGTGAAGTAACCTGGACCGTTTATCAACACAGCTTTAGAAAGAA 2051

QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTTCCTAAT 120
Db 2052 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTTCCTAAT 2111

QY 121 GATGTGGAAGTGGTGAAGACCAAGTTTCTACTCATGAGGGGTACATGATGATTTTGACA 180
Db 2112 GATGTGGAAGTGGTGAAGACCAAGTTTCTACTCATGAGGGGTACATGATGATTTTGACA 2171

QY 181 GCCCATCAGGGCCGGGTTGGTAATATTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGA 240
Db 2172 GCCCATCAGGGCCGGGTTGGTAATATTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGA 2231

QY 241 AAATTATCAGAAGATGAAGAAAACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGA 300
Db 2232 AAATTATCAGAAGATGAAGAAAACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGA 2291

QY 301 TGGGAATGCCTCAGGSTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATG 360
Db 2292 TGGGAATGCCTCAGGSTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATG 2351

QY 361 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGA 420
Db 2352 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGA 2411

QY 421 ACAGAGAAAAATGGAGGAAGAGCCTCTTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 480
Db 2412 ACAGAGAAAAATGGAGGAAGAGCCTCTTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 2471

QY 481 CAACAAACATAAGGTGCTTCAAGAAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTC 540
Db 2472 CAACAAACATAAGGTGCTTCAAGAAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTC 2531

QY 541 ACTCATGCTGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAGCAGCAACTGCTGCTTTGGAA 600
Db 2532 ACTCATGCTGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAGCAGCAACTGCTGCTTTGGAA 2591

QY 601 GAACAACTTAAGGTATTGGGAGATCGATGGGCAACACATCTGTAGATGGACAGAACCCGC 660
Db 2592 GAACAACTTAAGGTATTGGGAGATCGATGGGCAACACATCTGTAGATGGACAGAACCCGC 2651

QY 661 TGGGTTCTTTTACAAGACCAGCCTGACCTAGCTCCTGGACTGACCACCTATTGGAGCCTCT 720
Db 2652 TGGGTTCTTTTACAAGAC----- 2669

QY 721 CCTACTCAGACTGTTACTCTCTGGTGACACAACTGTGGTTACTAAGGAAACTGCCATCTCC 780
Db 2670 ----- 2669

QY 781 AAACCTAGAAATGCCATCTTCTCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 840
Db 2670 -----ACTCATAGATTACTGCAACAG 2690

QY 841 TTCCCCCTGGACCTGGAAAAAGTTTCTGCCTGGCTTACAGAAAGCTGAAACAACTGCCAAT 900
Db 2691 TTCCCCCTGGACCTGGAAAAAGTTTCTGCCTGGCTTACAGAAAGCTGAAACAACTGCCAAT 2750

QY 901 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTTAGAAGACTCCAAGGGAGTAAAGAG 960
Db 2751 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTTAGAAGACTCCAAGGGAGTAAAGAG 2810
```

2Y 961 CTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAC 1020
2b |||||
2Y 2811 CTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAC 2870
2b |||||
2Y 1021 AACCTGGATGAAAACAGCCAAAAATCCTGAGATCCCTGGAAAGTTCGGATGATGCAGTC 1080
2b |||||
2Y 2871 AACCTGGATGAAAACAGCCAAAAATCCTGAGATCCCTGGAAAGTTCGGATGATGCAGTC 2930
2b |||||
2Y 1081 CTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCT 1140
2b |||||
2Y 2931 CTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCT 2990
2b |||||
2Y 1141 CTCAACATTAGGTCCTTGGAAAGCCAGTTCTGACCAAGTGGAAAGCGTCTGCACCTTTCT 1200
2b |||||
2Y 2991 CTCAACATTAGGTCCTTGGAAAGCCAGTTCTGACCAAGTGGAAAGCGTCTGCACCTTTCT 3050
2b |||||
2Y 1201 CTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCT 1260
2b |||||
2Y 3051 CTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCT 3110
2b |||||
2Y 1261 ATTGGAGCGCACTTCCAGCAGTTTCAGAAAGCAGAACCATGTACATAGGCGCTTCAAGAGG 1320
2b |||||
2Y 3111 ATTGGAGCGCACTTCCAGCAGTTTCAGAAAGCAGAACCATGTACATAGGCGCTTCAAGAGG 3170
2b |||||
2Y 1321 GAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 1380
2b |||||
2Y 3171 GAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 3230
2b |||||
2Y 1381 ACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACAGAGGCCAGAGAGCTGCCTCCT 1440
2b |||||
2Y 3231 ACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACAGAGGCCAGAGAGCTGCCTCCT 3290
2b |||||
2Y 1441 GAGGAGAGGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACT 1500
2b |||||
2Y 3291 GAGGAGAGGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACT 3350
2b |||||
2Y 1501 GAGTGGGAAAAATTGAACCTCGACTCCGCTGACTGCGAGAGAAAAATAGATGAGACCCCTT 1560
2b |||||
2Y 3351 GAGTGGGAAAAATTGAACCTCGACTCCGCTGACTGCGAGAGAAAAATAGATGAGACCCCTT 3410
2b |||||
2Y 1561 GAAAGACTCCAGGAACCTTCAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAAGCT 1620
2b |||||
2Y 3411 GAAAGACTCCAGGAACCTTCAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAAGCT 3470
2b |||||
2Y 1621 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGAT 1680
2b |||||
2Y 3471 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGAT 3530
2b |||||
2Y 1681 CACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACGTGAGC 1740
2b |||||
2Y 3531 CACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACGTGAGC 3590
2b |||||
2Y 1741 CACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAAC 1800
2b |||||
2Y 3591 CACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAAC 3650
2b |||||
2Y 1801 CTCAGCACTCTGGAAGACCTGAAACCCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGAC 1860
2b |||||
2Y 3651 CTCAGCACTCTGGAAGACCTGAAACCCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGAC 3710
2b |||||
2Y 1861 CGAGTCAGGCAGCTGCATGAAGCCCAAGGGACTTTGGTCCAGCATCTCAGCACTTTCTT 1920
2b |||||
2Y 3711 CGAGTCAGGCAGCTGCATGAAGCCCAAGGGACTTTGGTCCAGCATCTCAGCACTTTCTT 3770
2b |||||
2Y 1921 TCCACGTCGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGSCAAAAAAGTGCCTACTAT 1980
2b |||||
2Y 3771 TCCACGTCGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGSCAAAAAAGTGCCTACTAT 3830
2b |||||
2Y 1981 ATCAACCCAGAGACTCAAACA 2001
2b |||||
2Y 3831 ATCAACCCAGAGACTCAAACA 3851
2b |||||

RESULT 8
US-09-845-416-2
; Sequence 2, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE THEREOF
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4182
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-2

Query Match 83.1%; Score 1662.8; DB 13; Length 4182;
Best Local Similarity 88.3%; Pred. No. 0;
Matches 1929; Conservative 0; Mismatches 72; Indels 183; Gaps 4;

QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA 60
Db 1000 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA 1059

QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAAT 120
Db 1060 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAAT 1119

QY 121 GATGTGGAAGTGGTGAAGACCAAGTTTCATCTACTCATGAGGGGTACATGATGGATTTGACA 180
Db 1120 GATGTGGAAGTGGTGAAGACCAAGTTTCATCTACTCATGAGGGGTACATGATGGATTTGACA 1179

QY 181 GCCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGA 240
Db 1180 GCCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGA 1239

QY 241 AATTATCAGAAAGATGAAGAACTGAAGTACAAGACAGATGAATCTCCTAATTTCAAGA 300
Db 1240 AATTATCAGAAAGATGAAGAACTGAAGTACAAGACAGATGAATCTCCTAATTTCAAGA 1299

QY 301 TGGGAATGCCTCAGGGTAGTCATGCAATGGAAAAACAAGCAATTTACATAGAGTTTAAATG 360
Db 1300 TGGGAATGCCTCAGGGTAGTCATGCAATGGAAAAACAAGCAATTTACATAGAGTTTAAATG 1359

QY 361 GATCTCCAGAATCAGAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAA 420
Db 1360 GATCTCCAGAATCAGAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAA 1419

QY 421 ACAAGGAAAAATGGAGGAAGACCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 480
Db 1420 ACAAGGAAAAATGGAGGAAGACCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 1479

QY 481 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGTCAAGTCAAGTCAATTTCTCTC 540
Db 1480 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGTCAAGTCAAGTCAATTTCTCTC 1539

QY 541 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 600
Db 1540 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 1599

QY 601 GAACAACTTAAGGTATTGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCGC 660
Db 1600 GAACAACTTAAGGTATTGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCGC 1659

QY 661 TGGGTTCTTTTACAAGACCCAGCCTGAC-----CTAGCTCCTGGACTGACCACT 708
Db 1660 TGGGTTCTTTTACAAGACATCCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTT 1719

QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTATACATAGAGTTTAAATG 360
Db 2057 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTATACATAGAGTTTAAATG 2116
QY 361 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGA 420
Db 2117 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGA 2176
QY 421 ACAAGGAAAAATGGAGGAAGAGCCTCTTGGACCTGTATCTTGAAGACCTAAACGCCAAGTA 480
Db 2177 ACAAGGAAAAATGGAGGAAGAGCCTCTTGGACCTGTATCTTGAAGACCTAAACGCCAAGTA 2236
QY 481 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTCTCTC 540
Db 2237 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGTCAGGGTCAATTCTCTC 2296
QY 541 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTTTGGAA 600
Db 2297 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTTTGGAA 2356
QY 601 GAACAACTTAAGGTATTTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGC 660
Db 2357 GAACAACTTAAGGTATTTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGC 2416
QY 661 TGGGTTCTTTTACAAGACCAGCCTGAC-----CTAGCTCTGGACTGACCACT 708
Db 2417 TGGGTTCTTTTACAAGACATCCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCCTT 2476
QY 709 ATTGGAGCCTCTCCTACTCAGACTGTGTACTCTGGTGACACA-----ACCTGTGGTT 759
Db 2477 TTTAGTGATGGCTTTCAGAAAAAGAGATGCAGTGAACAAGATTCACACAACCTGGCTTT 2536
QY 760 ACTAAGGAACTGCCATCTC-----779
Db 2537 AAAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTA 2596
QY 780 -----CAAAGTAAAGTCCATGGGCAAACTGTATTCACTCAAAAGATCTTCTTTCAACA 2656
Db 2597 GAAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAAGATCTTCTTTCAACA 2656
QY 802 TTGATGTTGGAG-----813
Db 2657 CTGAAGAATAAGTCAGTGACCCAGAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGT 2716
QY 814 -----GTACCTACTCATAGATTACTGCAA 837
Db 2717 TGGGATAATTTACTCCAAAACTTGAAAAAGAGTACAGCACAGACTCATAGATTACTGCAA 2776
QY 838 CAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCC 897
Db 2777 CAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCC 2836
QY 898 AATGTCCTACAGGATGCTACCCGTAAAGGAAAGGCTCCTAGAGACTCCAAGGGATAAA 957
Db 2837 AATGTCCTACAGGATGCTACCCGTAAAGGAAAGGCTCCTAGAGACTCCAAGGGATAAA 2896
QY 958 GAGCTGATGAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTAT 1017
Db 2897 GAGCTGATGAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTAT 2956
QY 1018 CACAACCTGGATGAAAAACAGCCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCA 1077
Db 2957 CACAACCTGGATGAAAAACAGCCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCA 3016
QY 1078 GTCCTGTTACAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTCGGAAAAAG 1137
Db 3017 GTCCTGTTACAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAG 3076
QY 1138 TCTCTCAACATTAGGTCCCATTTTGGAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTT 1197
Db 3077 TCTCTCAACATTAGGTCCCATTTTGGAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTT 3136

QY 1198 TCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCA 1257
Db 3137 TCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCA 3196
QY 1258 CCTATTGGAGGCGCACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAG 1317
Db 3197 CCTATTGGAGGCGCACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAG 3256
QY 1318 AGGGAATTGAAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTT 1377
Db 3257 AGGGAATTGAAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTT 3316
QY 1378 CTGACAGAGCAGCCTTTGGAAGGACTAGAGAAAACCTTACCAGGAGCCAGAGAGCTGCCT 1437
Db 3317 CTGACAGAGCAGCCTTTGGAAGGACTAGAGAAAACCTTACCAGGAGCCAGAGAGCTGCCT 3376
QY 1438 CCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAAT 1497
Db 3377 CCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAAT 3436
QY 1498 ACTGAGTGGGAAAAAATTGAACCTGCCTCGCTGACTGGCAGAGAAAAATAGATGAGACC 1557
Db 3437 ACTGAGTGGGAAAAAATTGAACCTGCCTGACTCGCTGACTGGCAGAGAAAAATAGATGAGACC 3496
QY 1558 CTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAA 1617
Db 3497 CTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAA 3556
QY 1618 GCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCTCATTTGACTCTCTCCAA 1677
Db 3557 GCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCTCATTTGACTCTCTCCAA 3616
QY 1678 GATCACCTCGAGAAAGTCAAGGCACTTCAGAGGAGAAATTGGCCCTCTGAAAGAGAACGTG 1737
Db 3617 GATCACCTCGAGAAAGTCAAGGCACTTCAGAGGAGAAATTGGCCCTCTGAAAGAGAACGTG 3676
QY 1738 AGCCACGTCAATGACCTTGCTGCGCAGCTTACCACCTTTGGGCAATTGAGCTCTCACCGTAT 1797
Db 3677 AGCCACGTCAATGACCTTGCTGCGCAGCTTACCACCTTTGGGCAATTGAGCTCTCACCGTAT 3736
QY 1798 AACCTCAGCACTCTTGGAAAGACCTGAACACCAAGATGGAAGCTTCTGCAGGTGGCCGTCGAG 1857
Db 3737 AACCTCAGCACTCTTGGAAAGACCTGAACACCAAGATGGAAGCTTCTGCAGGTGGCCGTCGAG 3796
QY 1858 GACCGAGTCAGGCAGCTGCATGAAGCCCAAGGCACTTTGGTCCAGCATCTCAGCACTTT 1917
Db 3797 GACCGAGTCAGGCAGCTGCATGAAGCCCAAGGCACTTTGGTCCAGCATCTCAGCACTTT 3856
QY 1918 CTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTTAC 1977
Db 3857 CTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTTAC 3916
QY 1978 TATATCAACCACGAGACTCAAAACA 2001
Db 3917 TATATCAACCACGAGACTCAAAACA 3940

RESULT 10
US-10-149-736-41
; Sequence 41, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96

```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; LENGTH: 5462
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-41

Query Match      64.1%; Score 1283; DB 13; Length 5462;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1302; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

QY 691 GCTCCTGGACTGACCACTATTGGAGCCCTCTCCTACTCAGACTGTTACTCTGGTGACACAA 750
Db 1547 GCTCCTGGACTGACCACTATTGGAGCCCTCTCCTACTCAGACTGTTACTCTGGTGACACAA 1606

QY 751 CCTGTGTTACTAAGGAAACTGCCATCTCCRAACTAGAAATGCCATCTTCTTGATGTTG 810
Db 1607 CCTGTGTTACTAAGGAAACTGCCATCTCCRAACTAGAAATGCCATCTTCTTGATGTTG 1666

QY 811 GAGGTACCTACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAAGTTCTTGCC 870
Db 1667 GAG-----CATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAAGTTCTTGCC 1717

QY 871 TGGCTTACAGAAAGCTGAAACAACTGCCAATGTCCTACAGGATGTCACCCGTAAGGAAAG 930
Db 1718 TGGCTTACAGAAAGCTGAAACAACTGCCAATGTCCTACAGGATGTCACCCGTAAGGAAAG 1777

QY 931 CTCCTAGAAAGACTCCAGGGAGTAAAGAGCTGTATGAAACAATGGCAAGACCTCCAAGGT 990
Db 1778 CTCCTAGAAAGACTCCAGGGAGTAAAGAGCTGTATGAAACAATGGCAAGACCTCCAAGGT 1837

QY 991 GAAATTGAAGCTCACACAGATGTTTATCAACACCTGGATGAAACAGCCCAAAAATCCTG 1050
Db 1838 GAAATTGAAGCTCACACAGATGTTTATCAACACCTGGATGAAACAGCCCAAAAATCCTG 1897

QY 1051 AGATCCCTGGAAGTTCGATGATGCAGTCCTCTGTTACAAAGACGTTTGGATAACATGAAC 1110
Db 1898 AGATCCCTGGAAGTTCGATGATGCAGTCCTCTGTTACAAAGACGTTTGGATAACATGAAC 1957

QY 1111 TTCAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAGTGCCCAATTTGGAAGCCAGT 1170
Db 1958 TTCAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAGTGCCCAATTTGGAAGCCAGT 2017

QY 1171 TCTGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTG 1230
Db 2018 TCTGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTG 2077

QY 1231 AAAGATGATGAATTAAAGCGGCAGGCACCTATTGGAGGCGACTTTCAGACAGTTTCAGAAG 1290
Db 2078 AAAGATGATGAATTAAAGCGGCAGGCACCTATTGGAGGCGACTTTCAGACAGTTTCAGAAG 2137

QY 1291 CAGAACGATGTACATAGGCGCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAAATCATG 1350
Db 2138 CAGAACGATGTACATAGGCGCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAAATCATG 2197

QY 1351 AGTACTCTTGAGACTGTACGAATATTTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAA 1410
Db 2198 AGTACTCTTGAGACTGTACGAATATTTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAA 2257

QY 1411 CTCTACAGGAGCCCGAGAGCTGCCTCCTGAGGAGAGAGCCCGAGAATGTCACTCGGCTT 1470
Db 2258 CTCTACAGGAGCCCGAGAGAGCTGCCTCCTGAGGAGAGAGCCCGAGAATGTCACTCGGCTT 2317

QY 1471 CTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACCTCGCT 1530
Db 2318 CTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACCTCGCT 2377

QY 1531 GACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTCAAGAGGCCACG 1590
Db 2378 GACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTCAAGAGGCCACG 2437
```

```

QY 1591 GATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCCGTG 1650
Db 2438 GATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCCGTG 2497

QY 1651 GCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGA 1710
Db 2498 GCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGA 2557

QY 1711 GAAATTGGCCCTCTGAAAGAGAACGCTGAGGCCACGTCATGACCTTGCTCGCCAGCTTACC 1770
Db 2558 GAAATTGGCCCTCTGAAAGAGAACGCTGAGGCCACGTCATGACCTTGCTCGCCAGCTTACC 2617

QY 1771 ACTTTGGGCAATCAGCTCTCACCCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGA 1830
Db 2618 ACTTTGGGCAATCAGCTCTCACCCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGA 2677

QY 1831 TGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCCACAG 1890
Db 2678 TGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCCACAG 2737

QY 1891 GACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGA 1950
Db 2738 GACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGA 2797

QY 1951 GCCATCTCGCCAAACAAAGTGCCCTACTATATATCAACCACGAGACTCAAAACA 2001
Db 2798 GCCATCTCGCCAAACAAAGTGCCCTACTATATATCAACCACGAGACTCAAAACA 2848

RESULT 11
US-10-149-736-42
; Sequence 42, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 8689
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-42

Query Match      59.1%; Score 1182.6; DB 13; Length 8689;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 813 GGTACCTACTATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAGTTTCTTGCCCTG 872
Db 2992 GGAAGAAACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAGTTTCTTGCCCTG 3051

QY 873 GCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGGCT 932
Db 3052 GCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGGCT 3111

QY 933 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 992
Db 3112 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 3171

QY 993 AATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAAACAGCCAAAAAATCCTGAG 1052
Db 3172 AATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAAACAGCCAAAAAATCCTGAG 3231
```


QY 1053 ATCCCTGGAAGSTCCGATGATGCAGTCCTGTGTACAAAGACGTTTGGATAACATGAACCTT 1112
DB |||||
QY 3232 ATCCCTGGAAGSTCCGATGATGCAGTCCTGTGTACAAAGACGTTTGGATAACATGAACCTT 3291
DB |||||
QY 1113 CAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAGGTCCCATTTTGGAGCCAGTTC 1172
DB |||||
QY 3292 CAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAGGTCCCATTTTGGAGCCAGTTC 3351
DB |||||
QY 1173 TGACCAGTGGAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGTGTGGCTACAGCTGAA 1232
DB |||||
QY 3352 TGACCAGTGGAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGTGTGGCTACAGCTGAA 3411
DB |||||
QY 1233 AGATGATGAATTAAGCCCGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAGCA 1292
DB |||||
QY 3412 AGATGATGAATTAAGCCCGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAGCA 3471
DB |||||
QY 1293 GAACGATGTACATAGGGCCCTTCAAGAGGGAAATTGAAAACCTAAAGAACCTGTAATCATGAG 1352
DB |||||
QY 3472 GAACGATGTACATAGGGCCCTTCAAGAGGGAAATTGAAAACCTAAAGAACCTGTAATCATGAG 3531
DB |||||
QY 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT 1412
DB |||||
QY 3532 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT 3591
DB |||||
QY 1413 CTACCAGGAGCCAGAGAGCTGCCCTCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCT 1472
DB |||||
QY 3592 CTACCAGGAGCCAGAGAGCTGCCCTCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCT 3651
DB |||||
QY 1473 ACGAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCCGCTGA 1532
DB |||||
QY 3652 ACGAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCCGCTGA 3711
DB |||||
QY 1533 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTCAAGAGGCCACGGA 1592
DB |||||
QY 3712 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTCAAGAGGCCACGGA 3771
DB |||||
QY 1593 TGAGCTGGACCTCAAGCTCGGCCAAGCTGAGGTGATCAAGGGATCTTGGCAGCCCGTGG 1652
DB |||||
QY 3772 TGAGCTGGACCTCAAGCTCGGCCAAGCTGAGGTGATCAAGGGATCTTGGCAGCCCGTGG 3831
DB |||||
QY 1653 CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 1712
DB |||||
QY 3832 CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 3891
DB |||||
QY 1713 AATTGCGCCTCTGAAGAGAACGCTGAGGCCACGTCAATGACCTTCTCTGCCAGCTTACCAC 1772
DB |||||
QY 3892 AATTGCGCCTCTGAAGAGAACGCTGAGGCCACGTCAATGACCTTCTCTGCCAGCTTACCAC 3951
DB |||||
QY 1773 TTTGGGCATTGAGTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCATG 1832
DB |||||
QY 3952 TTTGGGCATTGAGTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCATG 4011
DB |||||
QY 1833 GAAGCTTCTGAGGTGGCCGTGAGGACCGAGTCAGGACGCTGATGAAGCCACACGGA 1892
DB |||||
QY 4012 GAAGCTTCTGAGGTGGCCGTGAGGACCGAGTCAGGACGCTGATGAAGCCACACGGA 4071
DB |||||
QY 1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGTCTGTCTCAGGTCCCTGGGAGAGC 1952
DB |||||
QY 4072 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGTCTGTCTCAGGTCCCTGGGAGAGC 4131
DB |||||
QY 1953 CATCTGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACA 2001
DB |||||
QY 4132 CATCTGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACA 4180
DB |||||

RESULT 12
US-09-845-416-1
; Sequence 1, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE

; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11058
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-1

Query Match 59.1%; Score 1182.6; DB 13; Length 11058;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCTG 872
DB |||||
QY 8052 GGAAGAAACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCTG 8111
DB |||||
QY 873 GCTTACAGAGCTGAAACCAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGCT 932
DB |||||
QY 8112 GCTTACAGAGCTGAAACCAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGCT 8171
DB |||||
QY 933 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGSCAAGACCTCCAAGGTGA 992
DB |||||
QY 8172 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGSCAAGACCTCCAAGGTGA 8231
DB |||||
QY 993 AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAAATCCTGAG 1052
DB |||||
QY 8232 AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAAATCCTGAG 8291
DB |||||
QY 1053 ATCCCTGGAAGTTCGGATGATGCAGTCTCTTACAAAGACGTTTGGATAACATGAACCTT 1112
DB |||||
QY 8292 ATCCCTGGAAGTTCGGATGATGCAGTCTCTTACAAAGACGTTTGGATAACATGAACCTT 8351
DB |||||
QY 1113 CAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC 1172
DB |||||
QY 8352 CAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC 8411
DB |||||
QY 1173 TGACCAGTGGAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGTGTGGCTACAGCTGAA 1232
DB |||||
QY 8412 TGACCAGTGGAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGTGTGGCTACAGCTGAA 8471
DB |||||
QY 1233 AGATGATGAATTAAGCCCGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTTCAGAGCA 1292
DB |||||
QY 8472 AGATGATGAATTAAGCCCGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTTCAGAGCA 8531
DB |||||
QY 1293 GAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACTAAGAACCTGTATCATGAG 1352
DB |||||
QY 8532 GAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACTAAGAACCTGTATCATGAG 8591
DB |||||
QY 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT 1412
DB |||||
QY 8592 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT 8651
DB |||||
QY 1413 CTACCAGGAGCCAGAGAGCTGCCCTCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCT 1472
DB |||||
QY 8652 CTACCAGGAGCCAGAGAGCTGCCCTCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCT 8711
DB |||||
QY 1473 ACGAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCCGCTGA 1532
DB |||||
QY 8712 ACGAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCCGCTGA 8771
DB |||||
QY 1533 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTCAAGAGGCCACGGA 1592
DB |||||
QY 8772 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTCAAGAGGCCACGGA 8831
DB |||||
QY 1593 TGAGCTGGACCTCAAGCTCGGCCAAGCTGAGGTGATCAAGGGATCTTGGCAGCCCGTGG 1652
DB |||||
QY 8832 TGAGCTGGACCTCAAGCTCGGCCAAGCTGAGGTGATCAAGGGATCTTGGCAGCCCGTGG 8891
DB |||||

QY 1653 CGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 1712
Db 8892 CGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 8951
QY 1713 AATTGCGCCTCTGAAAGAGAACGTCGAGCCACGTCGAATGACCTTGCTGCGCCAGCTTACCAC 1772
Db 8952 AATTGCGCCTCTGAAAGAGAACGTCGAGCCACGTCGAATGACCTTGCTGCGCCAGCTTACCAC 9011
QY 1773 TTTGGGCATTTCAGCTCTCACCGGTATAAACCTCAGCACTCTGGAAGACCTGGAACACCCAGATG 1832
Db 9012 TTTGGGCATTTCAGCTCTCACCGGTATAAACCTCAGCACTCTGGAAGACCTGGAACACCCAGATG 9071
QY 1833 GAAGCTTCTGCAAGGTGGCGTCGAGGACCGAGTCGAGTCAGGCAGCTGCAATGAAAGCCCAAGGGA 1892
Db 9072 GAAGCTTCTGCAAGGTGGCGTCGAGGACCGAGTCGAGTCAGGCAGCTGCAATGAAAGCCCAAGGGA 9131
QY 1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCGTCCAGGTCCTGGGAGAGAGC 1952
Db 9132 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCGTCCAGGTCCTGGGAGAGAGC 9191
QY 1953 CATCTGCCAAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACA 2001
Db 9192 CATCTGCCAAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACA 9240

RESULT 13
US-10-149-736-44
; Sequence 44, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 11443
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic

Query Match 59.1%; Score 1182.6; DB 13; Length 11443;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCCCTCGACCTGGAAAAGTTTCTTGCCTG 872
Db 5746 GGAAGAAACTCATAGATTACTGCAACAGTTCCCTCGACCTGGAAAAGTTTCTTGCCTG 5805
QY 873 GCTTACAGAAAGCTGAAACAACTGCCAATGTCCTACAGATGCTACCCGTAAAGAAAGGCT 932
Db 5806 GCTTACAGAAAGCTGAAACAACTGCCAATGTCCTACAGATGCTACCCGTAAAGAAAGGCT 5865
QY 933 CCTAGAAAGACTCCAAGGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAAGGTGA 992
Db 5866 CCTAGAAAGACTCCAAGGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAAGGTGA 5925
QY 993 AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAGCCAAAATCCTGAG 1052
Db 5926 AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAGCCAAAATCCTGAG 5985
QY 1053 ATCCCTGGAAGGTTCCGATGATGAGTCTGTTTACAAGACGTTTGGATAACATGAACCTT 1112

Db 5986 ATCCCTGGAAGGTTCCGATGATGAGTCCTGTTTACAAAGACGTTTGGATAACATGAACCTT 6045
QY 1113 CAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCAATTTGGAAGCCAGTTC 1172
Db 6046 CAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCAATTTGGAAGCCAGTTC 6105
QY 1173 TGACCAGTGGAAAGGCTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 1232
Db 6106 TGACCAGTGGAAAGGCTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 6165
QY 1233 AGATGATGAATTAAGCCCGCAGGCACCTATTTGGAGGCGACCTTTCAGCAGTTCAGAAAGCA 1292
Db 6166 AGATGATGAATTAAGCCCGCAGGCACCTATTTGGAGGCGACCTTTCAGCAGTTCAGAAAGCA 6225
QY 1293 GAAAGATGTACATAGGGCCTTCAAGAGGGAATTTGAAAACTTAAAGAACCTGTAAATCATGAG 1352
Db 6226 GAAAGATGTACATAGGGCCTTCAAGAGGGAATTTGAAAACTTAAAGAACCTGTAAATCATGAG 6285
QY 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACT 1412
Db 6286 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACT 6345
QY 1413 CTACCAGGAGCCCAAGAGAGCTGCCCTCCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCT 1472
Db 6346 CTACCAGGAGCCCAAGAGAGAGCTGCCCTCCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCT 6405
QY 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACCTCCGCTGA 1532
Db 6406 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACCTCCGCTGA 6465
QY 1533 CTGGCAGAGAAAAATAGATGAGACCCCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACCGGA 1592
Db 6466 CTGGCAGAGAAAAATAGATGAGACCCCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACCGGA 6525
QY 1593 TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGTGATCAAGGGATCCTGGCAGCCCGTGGG 1652
Db 6526 TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGTGATCAAGGGATCCTGGCAGCCCGTGGG 6585
QY 1653 CGATCTCCTCATTTGACTCTCTCCAAGATCACTCGAGAAAGTCAAGGCACTTCGAGGAGA 1712
Db 6586 CGATCTCCTCATTTGACTCTCTCCAAGATCACTCGAGAAAGTCAAGGCACTTCGAGGAGA 6645
QY 1713 AATTGCGCCTCTGAAAGAGAACGTCGAGCCACGTCATGACCTTGTCTGCCAGCTTACCAC 1772
Db 6646 AATTGCGCCTCTGAAAGAGAACGTCGAGCCACGTCATGACCTTGTCTGCCAGCTTACCAC 6705
QY 1773 TTTGGGCATTTCAGCTCTCACCGTATAAACCTCAGCACTCTGGAAGACCTGAACACCCAGATG 1832
Db 6706 TTTGGGCATTTCAGCTCTCACCGTATAAACCTCAGCACTCTGGAAGACCTGAACACCCAGATG 6765
QY 1833 GAAGCTTCTGCAAGGTGGCGTCGAGGACCGAGTCGAGGACCGAGTCGATGAAGCCCAAGGGA 1892
Db 6766 GAAGCTTCTGCAAGGTGGCGTCGAGGACCGAGTCGAGGACCGAGTCGATGAAGCCCAAGGGA 6825
QY 1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCGTCCAGGTCCTGTTGGGAGAGAGC 1952
Db 6826 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCGTCCAGGTCCTGTTGGGAGAGAGC 6885
QY 1953 CATCTGCCAAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACA 2001
Db 6886 CATCTGCCAAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACA 6934

RESULT 14
US-10-149-736-47
; Sequence 47, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736

```

; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96

```

| | | | | |
|----------------------------|--------|---------------|--------|-------------------|
| Query Match | 59.1%; | Score 1182.6; | DB 13; | Length 12057; |
| Best Local Similarity | 99.7%; | Pred. No. 0; | | |
| Matches 1185; Conservative | 0; | Mismatches | 4; | Indels 0; Gaps 0; |

| | | | |
|----|------|---|------|
| QY | 813 | GGTACCTACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAGTTTCTTGGCTG | 872 |
| Db | 8260 | GGAGAAACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAGTTTCTTGGCTG | 8319 |
| QY | 873 | GCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGGCT | 932 |
| Db | 8320 | GCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGGCT | 8379 |
| QY | 933 | CCTAGAAGACTCCAAGGGAGTAAAGAGCTGTGATAAACAATGGCAAGACCTCCAGGTGA | 992 |
| Db | 8380 | CCTAGAAGACTCCAAGGGAGTAAAGAGCTGTGATAAACAATGGCAAGACCTCCAGGTGA | 8439 |
| QY | 993 | AATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAAAACAGCCAAAAATCCTGAG | 1052 |
| Db | 8440 | AATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAAAACAGCCAAAAATCCTGAG | 8499 |
| QY | 1053 | ATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTACAAAGACGTTTGGATAACATGAACTT | 1112 |
| Db | 8500 | ATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTACAAAGACGTTTGGATAACATGAACTT | 8559 |
| QY | 1113 | CAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCATTGTGAAGCCAGTTC | 1172 |
| Db | 8560 | CAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCATTGTGAAGCCAGTTC | 8619 |
| QY | 1173 | TGACCAGTGAAGGCTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGTGAA | 1232 |
| Db | 8620 | TGACCAGTGAAGGCTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGTGAA | 8679 |
| QY | 1233 | AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGCGAGTTCAGAAGCA | 1292 |
| Db | 8680 | AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGCGAGTTCAGAAGCA | 8739 |
| QY | 1293 | GAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTATCATGAG | 1352 |
| Db | 8740 | GAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTATCATGAG | 8799 |
| QY | 1353 | TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGAAGGACTAGAGAACT | 1412 |
| Db | 8800 | TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGAAGGACTAGAGAACT | 8859 |
| QY | 1413 | CTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAAATGTCACTCGGCTTCT | 1472 |
| Db | 8860 | CTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAAATGTCACTCGGCTTCT | 8919 |
| QY | 1473 | ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACCTCCGTGA | 1532 |
| Db | 8920 | ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACCTCCGTGA | 8979 |
| QY | 1533 | CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGA | 1592 |
| Db | 8980 | CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGA | 9039 |
| QY | 1593 | TGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCCGTGGG | 1652 |

| | | | |
|----|------|---|------|
| Qy | 1113 | CAAGTGGAGTGAACTTCGGA AAAAGTCTCTCAACATTAGGTCCCATTTTGAAGCCAGTTC | 1172 |
| Db | 8560 | CAAGTGGAGTGAACTTCGGA AAAAGTCTCTCAACATTAGGTCCCATTTTGAAGCCAGTTC | 8619 |
| Qy | 1173 | TGACCAGTGGAAAGCGTCTGCACCTTTCTCTGAGGAACTTCTGGTGGCTACAGCTGAA | 1232 |
| Db | 8620 | TGACCAGTGGAAAGCGTCTGCACCTTTCTCTGAGGAACTTCTGGTGGCTACAGCTGAA | 8679 |
| Qy | 1233 | AGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTTCAGCAGGTTTCAGAAAGCA | 1292 |
| Db | 8680 | AGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTTCAGCAGGTTTCAGAAAGCA | 8739 |
| Qy | 1293 | GAACGATGTACATAGGCGCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAAATCATGAG | 1352 |
| Db | 8740 | GAACGATGTACATAGGCGCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAAATCATGAG | 8799 |
| Qy | 1353 | TACTCTTGAGACTGTACGAATATTCTTGACAGAGCAGCCCTTTGGAGGACTAGAGAAACT | 1412 |
| Db | 8800 | TACTCTTGAGACTGTACGAATATTCTTGACAGAGCAGCCCTTTGGAGGACTAGAGAAACT | 8859 |
| Qy | 1413 | CTACCAGGAGCCGACAGAGCTGCCTCCTGAGGAGAGAGCCGAGAAATGTCACTCGGCTTCT | 1472 |
| Db | 8860 | CTACCAGGAGCCGACAGAGCTGCCTCCTGAGGAGAGAGCCGAGAAATGTCACTCGGCTTCT | 8919 |
| Qy | 1473 | ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACTGCACTCCGCTGA | 1532 |
| Db | 8920 | ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACTGCACTCCGCTGA | 8979 |
| Qy | 1533 | CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTCAAGAGGCCACCGGA | 1592 |
| Db | 8980 | CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTCAAGAGGCCACCGGA | 9039 |
| Qy | 1593 | TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGTGATCAAGGGATCCTGGCAGCCCGTGSG | 1652 |
| Db | 9040 | TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGTGATCAAGGGATCCTGGCAGCCCGTGSG | 9099 |
| Qy | 1653 | CGATCTCCTCATTTGACTCTCTCCAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGGAGA | 1712 |
| Db | 9100 | CGATCTCCTCATTTGACTCTCTCCAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGGAGA | 9159 |
| Qy | 1713 | AATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC | 1772 |
| Db | 9160 | AATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC | 9219 |
| Qy | 1773 | TTTGGGCATTACGCTCTCACCGGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATG | 1832 |
| Db | 9220 | TTTGGGCATTACGCTCTCACCGGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATG | 9279 |
| Qy | 1833 | GAAAGCTTCTGAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGGA | 1892 |
| Db | 9280 | GAAAGCTTCTGAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGGA | 9339 |
| Qy | 1893 | CTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGTCTGTCCAGGTCCTGGGAGAGAGC | 1952 |
| Db | 9340 | CTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGTCTGTCCAGGTCCTGGGAGAGAGC | 9399 |
| Qy | 1953 | CATCTGCCCAAAACAAAGTGCCCTACTATATCAACCAACGAGACTCAAACA | 2001 |
| Db | 9400 | CATCTGCCCAAAACAAAGTGCCCTACTATATCAACCAACGAGACTCAAACA | 9448 |

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 11:11:40 ; Search time 123.697 Seconds
(without alignments)
7140.092 Million cell updates/sec

```
Title: US-09-845-416-6_COPY_1000_3000
Perfect score: 2001
Sequence: 1 ggcagtccattgatggagag.....tcaaccacgagactcaaac 2001
```

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 s

```
Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCtUS COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|---------|----|---------------------|--------------------|
| 1 | 1182.6 | 59.1 | 5952 | 4 | US-09-687-875A-1 | Sequence 1, Appli |
| 2 | 1182.6 | 59.1 | 13977 | 4 | US-09-484-970B-60 | Sequence 60, Appl |
| C 3 | 1004 | 50.2 | 19307 | 3 | US-08-836-022A-10 | Sequence 10, Appl |
| C 4 | 1004 | 50.2 | 19307 | 3 | US-09-427-048A-10 | Sequence 10, Appl |
| 5 | 397 | 19.8 | 6045 | 4 | US-09-091-501B-7 | Sequence 7, Appli |
| 6 | 397 | 19.8 | 10320 | 4 | US-09-091-501B-9 | Sequence 9, Appli |
| 7 | 79.4 | 4.0 | 200 | 4 | US-09-091-501B-5 | Sequence 5, Appli |
| 8 | 78.6 | 3.9 | 200 | 4 | US-09-091-501B-4 | Sequence 4, Appli |
| 9 | 78.6 | 3.9 | 200 | 4 | US-09-091-501B-6 | Sequence 6, Appli |
| C 10 | 76.6 | 3.8 | 7218 | 1 | US-08-232-463-14 | Sequence 14, Appl |
| 11 | 63.6 | 3.2 | 238 | 4 | US-09-687-875A-13 | Sequence 13, Appl |
| C 12 | 44 | 2.2 | 1230025 | 4 | US-09-198-452A-1 | Sequence 1, Appli |
| 13 | 43.4 | 2.2 | 1179 | 4 | US-09-107-532A-1186 | Sequence 1186, Ap |
| 14 | 42.8 | 2.1 | 1690 | 4 | US-09-620-312D-69 | Sequence 69, Appl |
| 15 | 42.8 | 2.1 | 7812 | 3 | US-09-368-590-1 | Sequence 1, Appli |
| 16 | 40.4 | 2.0 | 2223 | 1 | US-08-257-073-4 | Sequence 4, Appli |
| 17 | 39.2 | 2.0 | 16995 | 4 | US-08-961-527-82 | Sequence 82, Appl |
| 18 | 38.6 | 1.9 | 1751 | 4 | US-09-620-312D-847 | Sequence 847, App |
| 19 | 38.6 | 1.9 | 1995 | 1 | US-08-425-069-3 | Sequence 3, Appli |
| 20 | 38.6 | 1.9 | 1995 | 2 | US-08-317-844B-3 | Sequence 3, Appli |
| 21 | 38.4 | 1.9 | 7672 | 4 | US-09-220-132-24 | Sequence 24, Appl |
| 22 | 38.2 | 1.9 | 1131 | 6 | 5180810-3 | Patent No. 5180810 |
| 23 | 38.2 | 1.9 | 1784 | 6 | 5180810-2 | Patent No. 5180810 |
| C 24 | 38 | 1.9 | 1394 | 4 | US-09-247-155-76 | Sequence 76, Appl |
| 25 | 36.8 | 1.8 | 1886 | 6 | 5210183-1 | Patent No. 5210183 |
| 26 | 36.6 | 1.8 | 1845 | 4 | US-08-887-534A-22 | Sequence 22, Appl |
| 27 | 36.6 | 1.8 | 1845 | 4 | US-09-527-431-22 | Sequence 22, Appl |

ALIGNMENTS

```

RESULT 1
US-03-687-875A-1
; Sequence 1, Application US/09687875A
; Patent No. 6544786
; GENERAL INFORMATION:
; APPLICANT: Xiao, Xiao
; APPLICANT: Liu, Paul
; TITLE OF INVENTION: METHOD AND VECTOR
; FILE REFERENCE: 00792
; CURRENT APPLICATION NUMBER: US/09/687875A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/158,866
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 5952
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2897)..(2898)
; OTHER INFORMATION: S4 junction site
; NAME/KEY: misc feature
; LOCATION: (3198)..(3199)
; OTHER INFORMATION: S2 junction site
; US-03-687-875A-1

```

| | | | | |
|----------------------------|--------|---------------|-----------|--------------|
| Query Match | 59.1%; | Score 1182.6; | DB 4; | Length 5952; |
| Best Local Similarity | 99.7%; | Pred. No. 0; | | |
| Matches 1185; Conservative | 0; | Mismatches 4; | Indels 0; | Gaps 0; |

| | | | |
|----|------|--|------|
| QY | 813 | GGTACCTACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAGTTTCTTGCCCTG | 872 |
| Db | 2946 | GGAGAAACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAGTTTCTTGCCCTG | 3005 |
| QY | 873 | GCTTACAGAAAGTGTGAAAACAACCTGCCAATGTCTCTACAGGATGCTACCCCGTAAGGAAAGGCT | 932 |
| Db | 3006 | GCTTACAGAAAGTGTGAAAACAACCTGCCAATGTCTCTACAGGATGCTACCCCGTAAGGAAAGGCT | 3065 |
| QY | 933 | CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAAACAATGGCAAGACCTCCAAAGGTGA | 992 |
| Db | 3066 | CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAAACAATGGCAAGACCTCCAAAGGTGA | 3125 |
| QY | 993 | AATTGAAGCTCACACAGATGTTTTATCAACAACCTGGATGAAAAACAGCCAAAAAATCCTGAG | 1052 |
| Db | 3126 | AATTGAAGCTCACACAGATGTTTTATCAACAACCTGGATGAAAAACAGCCAAAAAATCCTGAG | 3185 |
| QY | 1053 | ATCCCTGGAAGTTCCTCGATGATGCAGTCTCTGTACAAAGACGTTTGGATAACATGAACCTT | 1112 |

Db 3186 ATCCCTGGAAGGTTCCGATGATGTCAGTCTCTGTTACAAAGACGTTTGGATAACATGAACACTT 3245
QY 1113 CAAGTGAGTGAACTTCGGAAGAAAGTCTCTCAACATTAAGTCCCATTTTGGAGCCAGTTT 1172
Db 3246 CAAGTGAGTGAACTTCGGAAGAAAGTCTCTCAACATTAAGTCCCATTTTGGAGCCAGTTT 3305
QY 1173 TGACCACTGGAAGCGCTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 1232
Db 3306 TGACCACTGGAAGCGCTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 3365
QY 1233 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAAGAAGCA 1292
Db 3366 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAAGAAGCA 3425
QY 1293 GAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACCTAAAGAACCTGTATCATGAG 1352
Db 3426 GAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACCTAAAGAACCTGTATCATGAG 3485
QY 1353 TACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACT 1412
Db 3486 TACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACT 3545
QY 1413 CTACCAGAGCCCGAGAGAGCTGCCCTCTGAGGAGAGAGCCCGAGAAATGTCACTCGGCTTCT 1472
Db 3546 CTACCAGAGCCCGAGAGAGCTGCCCTCTGAGGAGAGAGCCCGAGAAATGTCACTCGGCTTCT 3605
QY 1473 ACGAAGCAGCGCTGAGGAGGTCAATCTAGTGGGAAAATTGAACCTGCACCTCGCTGA 1532
Db 3606 ACGAAGCAGCGCTGAGGAGGTCAATCTAGTGGGAAAATTGAACCTGCACCTCGCTGA 3665
QY 1533 CTGGCAGAGAAAATAGATGAGACCTTGAAAGACTCCAGAACTTCAAGAGGCCACGGA 1592
Db 3666 CTGGCAGAGAAAATAGATGAGACCTTGAAAGACTCCAGAACTTCAAGAGGCCACGGA 3725
QY 1593 TGAGCTGGACCTCAAGCTGCGCCAGCTGAGGTGATCAAGGGATCCTGGCAGCCGCTGGG 1652
Db 3726 TGAGCTGGACCTCAAGCTGCGCCAGCTGAGGTGATCAAGGGATCCTGGCAGCCGCTGGG 3785
QY 1653 CGATCTCCTCACTGACTCTCTCCAGATCACCTCGAGAAATCAAGGCACCTTCGAGGAGA 1712
Db 3786 CGATCTCCTCACTGACTCTCTCCAGATCACCTCGAGAAATCAAGGCACCTTCGAGGAGA 3845
QY 1713 AATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTCTCGCCAGCTTACCAC 1772
Db 3846 AATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTCTCGCCAGCTTACCAC 3905
QY 1773 TTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATG 1832
Db 3906 TTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATG 3965
QY 1833 GAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGA 1892
Db 3966 GAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGA 4025
QY 1893 CTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTGTCCAGGTCCTTGGGAGAGAGC 1952
Db 4026 CTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTGTCCAGGTCCTTGGGAGAGAGC 4085
QY 1953 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACA 2001
Db 4086 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACA 4134

RESULT 2

US-09-484-970B-60
; Sequence 60, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmuth, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US

; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 60
; LENGTH: 13977
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6426186 229357.11CB1
; NAME/KEY: unsure
; LOCATION: 11721-11761, 12294, 13969
; OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-60

Query Match 59.1%; Score 1182.6; DB 4; Length 13977;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTG 872
Db 8260 GGAAGAAACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTG 8319
QY 873 GCTTACAGAAGCTGAAACAACCTGCCAATGTCTACAGGATGCTACCCGTAAAGGAAAGGCT 932
Db 8320 GCTTACAGAAGCTGAAACAACCTGCCAATGTCTACAGGATGCTACCCGTAAAGGAAAGGCT 8379
QY 933 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAGGTTGA 992
Db 8380 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAGGTTGA 8439
QY 993 AATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAAACAGCCAAAATCCTGAG 1052
Db 8440 AATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAAACAGCCAAAATCCTGAG 8499
QY 1053 ATCCCTGGAAGGTTCCGATGATGTCAGTCTCTGTTACAAAGACGTTTGGATAACATGAACCT 1112
Db 8500 ATCCCTGGAAGGTTCCGATGATGTCAGTCTCTGTTACAAAGACGTTTGGATAACATGAACCT 8559
QY 1113 CAAGTGGAGTGAACTTCGGAAGAAAGTCTCTCAACATTAGGTCCCATTGGAAGCCAGTTT 1172
Db 8560 CAAGTGGAGTGAACTTCGGAAGAAAGTCTCTCAACATTAGGTCCCATTGGAAGCCAGTTT 8619
QY 1173 TGACCACTGGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 1232
Db 8620 TGACCACTGGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 8679
QY 1233 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAAGAGCA 1292
Db 8680 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAAGAGCA 8739
QY 1293 GAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACCTAAAGAACTGTAATCATGAG 1352
Db 8740 GAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACCTAAAGAACTGTAATCATGAG 8799
QY 1353 TACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT 1412
Db 8800 TACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT 8859
QY 1413 CTACCAGAGCCCGAGAGAGCTGCCCTCTGAGGAGAGAGCCCGAGAAATGTCACTCGGCTTCT 1472
Db 8860 CTACCAGAGCCCGAGAGAGCTGCCCTCTGAGGAGAGAGCCCGAGAAATGTCACTCGGCTTCT 8919
QY 1473 ACGAAAGCAGGCTGAGGAGGTCATTAATACTAGTGGGAAAATTAAGAACCTGCACCTCCGCTGA 1532
Db 8920 ACGAAAGCAGGCTGAGGAGGTCATTAATACTAGTGGGAAAATTAAGAACCTGCACCTCCGCTGA 8979
QY 1533 CTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGA 1592
Db 8980 CTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGA 9039
QY 1593 TGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGAATCCTGGCAGCCCCGTGG 1652

db 9040 TGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGG 9099
2y 1653 CGATCTCCTCATTGACTCTCTCCAAAGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGA 1712
db 9100 CGATCTCCTCATTGACTCTCTCCAAAGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGA 9159
2y 1713 AATTGGCCCTCTGAAAGAGAACGCTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCAC 1772
db 9160 AATTGGCCCTCTGAAAGAGAACGCTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCAC 9219
2y 1773 TTTGGCATTGAGCTCTCACCGTATACCTCAGCACTCTGGAAGACCTGAAACACAGATG 1832
db 9220 TTTGGCATTGAGCTCTCACCGTATACCTCAGCACTCTGGAAGACCTGAAACACAGATG 9279
2y 1833 GAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGA 1892
db 9280 GAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGA 9339
2y 1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCCTGCCAGGTCCCTGGGAGAGAGC 1952
db 9340 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCCTGCCAGGTCCCTGGGAGAGAGC 9399
1953 CATCTCGCCAAAACAAAGTGCCTTACTATATCAACCACGAGACTCAAACA 2001
db 9400 CATCTCGCCAAAACAAAGTGCCTTACTATATCAACCACGAGACTCAAACA 9448

RESULT 3

JS-08-836-022A-10/c
; Sequence 10, Application US/08836022A
; Patent No. 6001557
; GENERAL INFORMATION:
; APPLICANT: Trustees of the University of Pennsylvania
; APPLICANT: Wilson, James M.
; APPLICANT: Fisher, Krishna J.
; APPLICANT: Chen, Shu-Jen
; APPLICANT: Weitzman, Matthew
; TITLE OF INVENTION: Improved Adenovirus Virus and
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P O Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,022A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/331,381
; FILING DATE: 28-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GNVNPN.008PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19307 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA

US-08-836-022A-10
Query Match 50.2%; Score 1004; DB.3; Length 19307;
Best Local Similarity 90.3%; Pred. No. 6.1e-302;
Matches 1073; Conservative 0; Mismatches 115; Indels 0; Gaps 0;
Qy 813 GGTACCTACTCATAGATTACTGCAACAGTTCCCTCCCTGGACCTGGAAAAGTTTCTTCCTG 872
Db 6434 GGAAGAAACTCATAGATTACTGCAAGCAGTTCCCTCTGGACCTGGAGAAAGTTTCTTCCTG 6375
Qy 873 GCTTACAGAAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGCT 932
Db 6374 GATTACGGAAGCAGAAACAACTGCCAATGTCTCTACAGGACGCTTCCCGTAAGGAGAGCT 6315
Qy 933 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAACAAATGGCAAGACCTCCAAGGTGA 992
Db 6314 CCTAGAAGACTCCAGGGAGTCAGAGAGCTGATGAACCAATGGCAAGATCTCCAAGGAGA 6255
Qy 993 AATTGAAGCTCACACAGATGTTTATCAACACTGGAATGAACCAAGCAAGCAAAATCCTGAG 1052
Db 6254 AATTGAAGCTCACACAGATATCTATCACAATCTTGTATGAATAATGGCCAAAATCCTGAG 6195
Qy 1053 ATCCCTGGAAGGTTCCGATGATGACGTCTCTTACAAAGACGTTTGGATAACATGAACCTT 1112
Db 6194 ATCCCTGGAAGGTTCCGATGAAGCACCCCTGTTTACAAAGACGTTTGGATAACATGAATTT 6135
Qy 1113 CAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTGGGAAGCAAGTTT 1172
Db 6134 CAAGTGGAGTGAACCTTCAGAAAAAGTCTCTCAACATTAGGTCCCATTGGGAAGCAAGTTT 6075
Qy 1173 TGACCAAGTGAAGCGTCTGCACTTTCTCTGCAAGAACTTCTGGTGTGGCTACAGCTGAA 1232
Db 6074 TGACCAAGTGAAGCGTTTGCATCTTCTCTCAGGAACCTTCTTGTGGCTACAGCTGAA 6015
Qy 1233 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGACCTTCCAGCAGTTTCAGAAAGCA 1292
Db 6014 AGATGATGAACCTGAGCCGTCAGGCACCCATCGGTGGTGATTTCCAGCAGTTTCAGAAAGCA 5955
Qy 1293 GAACGATGTACATAGGGCCCTCAAGAGGGAATTTGAATACTAAAGAACCTGTATCATGAG 1352
Db 5954 GAATGATATACATAGGGCCCTCAAGAGGGAATTTGAATACTAAAGAACCTGTATCATGAG 5895
Qy 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT 1412
Db 5894 TACTCTTGAGACTGTGAGATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAACT 5835
Qy 1413 CTACCAAGAGCCCGAGAGAGCTGCTCTCTGAGGAGAGAGCCCAAGAACTGTCACTCGGCTTCT 1472
Db 5834 CTACCAAGAGCCCGAGAGAACTGCTCTCTGAAAGAAAGGTCAGAACTTTCAGGAAGCTGCCGA 5775
Qy 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAATAATGAACTGCACTCCGCTGA 1532
Db 5774 ACGAAAGCAGGCTGAAGAGGTCAACGCTGAATGGGACAAATTTGAACCTGCGCTCAGCTGA 5715
Qy 1533 CTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACCGGA 1592
Db 5714 TTGGCAGAGAAAAATAGATGAGTGAAGTCTTTGAAAGACTCCAGGAACTTTCAGGAAGCTGCCGA 5655
Qy 1593 TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 1652
Db 5654 TGAACCTGGACCTCAAGTTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCAGTGGG 5595
Qy 1653 CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 1712
Db 5594 GGATCTCCTCATTGACTCTCTGCAAGATCACCTTGAATAAGTCAAGGCACCTTCGCGGAGA 5535
Qy 1713 AATTGGCCCTCTGAAAGAGAACGTCAGCCACGTCATGACCTTGCTCGCCAGCTTACCAC 1772
Db 5534 AATTGCACCTCTTAAAGAGAAATGTCAATCGTGTCAATGACCTTGCAATCAGCTGACCAC 5475
Qy 1773 TTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACAGATG 1832
Db 5474 ACTGGGCATTGAGCTCTCACCTTATAACCTCAGCACTTTTGGAAAGATCTGAATACCAGATG 5415

RESULT 5
US-09-091-501B-7
; Sequence 7, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 6045
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(6037)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric
; NAME/KEY: misc_feature
; LOCATION: (724)..(758)
; OTHER INFORMATION: Precise residue is left open

Query Match 19.8%; Score 397; DB 4; Length 6045;
Best Local Similarity 59.3%; Pred. No. 6.8e-113;
Matches 698; Conservative 0; Mismatches 470; Indels 9; Gaps 1;
QY 833 TGCAACAGTCCCGCTGGACCTGGAAAGTTCTTGGCTTACAGAAAGCTGAAACAA 892
Db 3069 TGCAGGCCCTCTCGAGAGATCTGGAAACTTCTGAAAGTGGATCCAAAGAGCAGAGACCA 3128
QY 893 CTGCCAATGCTTACAGGATGCTACCCGTAAGGAAAGGCTCCAGAAAGACTCCAAGGGAG 952
Db 3129 CAGTGAATGCTTGTGGATGCTCTCATCGGGAGAAATGCTCTTCAGGATAGTATCTTGG 3188
QY 953 TAAAGAGCTGATGAAACAATGGCAAGACCTCCAAAGGTGAAATTAAGCTCACACAGATG 1012
Db 3189 CCAGGGAACCTCAACAGCAGATGCAGGACATCCAGGCAGAAATTTGATGCCACAATGACA 3248
QY 1013 TTATCACAACTGGATGAAACAGCCAAATAATCCTGAGATCCCTGGAAGTTCCGATG 1072
Db 3249 TATTAAAGCATTCAGCGAAACAGGCAGAGATGGTAAAGCTTTGGGAAATTTCTGAAG 3308
QY 1073 ATGCAGTCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTCGGA 1132
Db 3309 AGGCTACTATGCTTACATCGACTGGATGATGATGAACCAAGATGGAATGACTTAAAG 3368
QY 1133 AAAAGTCTCTCAACATTAGGTCCCATTTTGGAGCCAGTTCTGACCAGTGAAGCGTCTGC 1192
Db 3369 CAAAATCTGTAGCATCAGGGCCCCATTTGGAGCCAGCGCTGAGAACTGGAACAGGTTGC 3428
QY 1193 ACCTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGC 1252
Db 3429 TGATGCTCTTAGAAGAACTGATCAAAATGGCTGAATATGAAAGATGAAGAGCTTAAGAAAC 3488
QY 1253 AGCCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCT 1312
Db 3489 AAATGCCTATTGGAGGAGATGTTCCAGCCTTACAGCTCCAGTATGACCATTGTAAGGCC 3548

QY 1313 TCAAGAGGGAATTTGAAACCTAAGAACCTGTAAATCATGAGTACTCTTTGAGACTGTACGAA 1372
Db 3549 TGAGACGGGAGTTAAAGGAGAGAAATATTCTGTCTCTGAATGCTGTGACACGCGCCGAG 3608
QY 1373 TATTTCTGACAGAGCAGCCTTT-----GGAAGGACTAGAGAAACTCTACCAGGAGC 1423
Db 3609 TTTTCTTGGCTGATCAGCCAATTGAGGCCCTTGAAGAGGCCAAAGAAACCTTACAATCAA 3668
QY 1424 CCAGAGAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTCATCTCGCTTCTACGAAAGCAGG 1483
Db 3669 AACAGAAATTAATCTCTGAGGAGAGAGCCCAAAAGATTGCCAAAGCCATCGCAAAACAGT 3728
QY 1484 CTGAGGAGGTCAATACTAGTGGGAAAAAATTTGAACCTGCACCTCCGCTGACTGGCAGAGAA 1543
Db 3729 CTTCTGAAGTCAAGAAATAATGGGAAAGTCTAAATGCTGTAACTAGCAATTTGGCAAAAGC 3788
QY 1544 AAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACCGGATGAGCTGGACC 1603
Db 3789 AAGTGGACAGGCATTTGGAGAACTCAGAGACCTGCAGGGAGCTATGGATGACCTGGACG 3848
QY 1604 TCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCCTTGGCAGCCGCTGGCGATCTCCTCA 1663
Db 3849 CTGACATGAAGAGGCAGAGTCCGTCGGAATGGCTGGAAGCCGCTGGGAGACTTACTCA 3908
QY 1664 TTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCACCTTCAGAGGAGAAATTTGCGCCTC 1723
Db 3909 TTGACTCGCTGCAGGATCACATTTGAAAAAATCATGGCATTTAGAGAGAAATTTGCCACCA 3968
QY 1724 TGAAGAGAGACGTGAGCCACGTCATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTC 1783
Db 3969 TCAACTTTAAAGTTAAACGGTGAATGATTTATCCAGTCAGCTGTCTCCACTTGACCTGC 4028
QY 1784 AGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACAGATGGAAGCTTCTTGC 1843
Db 4029 ATCCCTCTCTAAAGATGTCTCGCCAGTAGATGACCTTATATATGCGATGGAACCTTTTAC 4088
QY 1844 AGGTGGCCGTCGAGGACCGAGTCAGCAGCTGAGCAGCTGCATGAAGCCACAGGGACTTTGGTCCAG 1903
Db 4089 AGGTTTCTGTGGATGATCGCCTTAAACAGCTTCAGGAAGCCACAGAGATTTTGGACCAT 4148
QY 1904 CATCTCAGCACTTTCTTTCCACGCTGTCTCCAGGTCCTTGGAGAGAGCCATCTCGCCAA 1963
Db 4149 CCTCTCAGCACTTTCTCTCTACGTCAGTCCAGCTGCCGTCAGGATGCAATTTTCCATA 4208
QY 1964 ACAAAGTGCCTACTATATATCAACACAGAGACTCAAAAC 2000
Db 4209 ATAAAGTGCCTATTACATCAACCATCAAAACACAGAC 4245

RESULT 6
US-09-091-501B-9
; Sequence 9, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 10320

| | | | | | | | | | |
|--|------|---|------|--|--|--|--|--|--|
| ; TYPE: DNA | | | | | | | | | |
| ; ORGANISM: Artificial Sequence | | | | | | | | | |
| ; FEATURE: | | | | | | | | | |
| ; NAME/KEY: CDS | | | | | | | | | |
| ; LOCATION: (11)..(10312) | | | | | | | | | |
| ; FEATURE: | | | | | | | | | |
| ; OTHER INFORMATION: Description of Artificial Sequence: Full length | | | | | | | | | |
| ; OTHER INFORMATION: utrophin construct | | | | | | | | | |
| ; FEATURE: | | | | | | | | | |
| ; NAME/KEY: misc.feature | | | | | | | | | |
| ; LOCATION: (724)..(758) | | | | | | | | | |
| ; OTHER INFORMATION: Precise residue is left open | | | | | | | | | |
| US-09-091-501B-9 | | | | | | | | | |
| Query Match | | | | | | | | | |
| Best Local Similarity 19.8%; Score 397; DB 4; Length 10320; | | | | | | | | | |
| Matches 698; Conservative 0; Mismatches 470; Indels 9; Gaps 1; | | | | | | | | | |
| QY | 833 | TGCAACAGTTCCCTGGACCTGGAAAAGTTTCTTGCTTGCTTACAGAAGCTGAAACAA | 892 | | | | | | |
| DB | 7344 | TGCAGGCCCTCTCGCAGAGATCTGGAAAACCTTCTGAAGTGGATCCAAAGAAGCAGAGACCA | 7403 | | | | | | |
| QY | 893 | CTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAAGGCTCCTAGAAGACTCCAAGGGAG | 952 | | | | | | |
| DB | 7404 | CAGTGAATGTGCTTGTGGATGCCCTCTCATCGGAGAAATGCTCTTCAGGATAGTATCTTGG | 7463 | | | | | | |
| QY | 953 | TAAAGAGCTGATGAACAATGGCAAGACCTTCAAGGTGAAATTGAAGTCTCACACAGATG | 1012 | | | | | | |
| DB | 7464 | CCAGGGAACCTCAACACAGCAGATGCAGGACATCCAGGCAGAAATTGATGCCCAATGACA | 7523 | | | | | | |
| QY | 1013 | TTTATCACAACTGGATGAAAACAGCCCCAAAATCTTGAGATCCCTGGAAAGTTCGGATG | 1072 | | | | | | |
| DB | 7524 | TATTTAAAAGCATTCACGGAACACAGGCAGAAAGATGGTAAAAGCTTTGGGAAATTTCTGAAG | 7583 | | | | | | |
| QY | 1073 | ATGCAGTCTCTTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGA | 1132 | | | | | | |
| DB | 7584 | AGGCTACTATGCTTCAACATCGACTGGATGATATGAACCAAGATGGAATGACTTAAAG | 7643 | | | | | | |
| QY | 1133 | AAAAGTCTCTCAACATTAGGTCCCATTTTGGAAAGCCAGTTCTTGACCAAGTGAAGCGTCTGC | 1192 | | | | | | |
| DB | 7644 | CAAAATCTGCTAGCATCAGGGCCCCATTTTGGAGGCCAGCGCTGAGAAGTGGAACAGGTTGC | 7703 | | | | | | |
| QY | 1193 | ACCTTTCTCTGCAGGAATCTTGGTGTGGCTACAGTGAAGATGATGAATTAAGCCGCGC | 1252 | | | | | | |
| DB | 7704 | TGATGTCTTTAGAAGAACTGATCAAATGGCTGAATATGAAGAATGAAGAGCTTAAGAAAC | 7763 | | | | | | |
| QY | 1253 | AGGCACCTATTGGAGGCGACTTTCAGCAGTTTCAGAAGCAGAACGATGTACATPAGGCGCT | 1312 | | | | | | |
| DB | 7764 | AAATGCCCTATTGGAGGAGATGTTCCAGCCTTACAGCTCCAGTATGACCATTTGTAAGGCCCC | 7823 | | | | | | |
| QY | 1313 | TCAAGAGGGAATTGAAAACATAAGAACCTGTATCATCATGAGTACTCTTGAGACTGTACGAA | 1372 | | | | | | |
| DB | 7824 | TGAGACGGGAGTTAAAGGAGAAAGAAATATTCTGTCTCTGAATGCTGTGACACCGCCGAG | 7883 | | | | | | |
| QY | 1373 | TATTTCTGACAGAGCAGCTTT-----GGAAGSACTAGAGAACTCTACAGGAGC | 1423 | | | | | | |
| DB | 7884 | TTTTCTGGCTGATCAGCCCAATTGAGGCCCTGAAGAGCCCAAGAAACCTTACAATCAA | 7943 | | | | | | |
| QY | 1424 | CCAGAGAGCTGCCCTCTGAGGAGAGAGCCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGG | 1483 | | | | | | |
| DB | 7944 | AAACAGAAATTAACCTCCTGAGGAGAGAGCCCCAAAAGATTGCCAAAGCCATGCCGAAACAGT | 8003 | | | | | | |
| QY | 1484 | CTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTCGCACTCCGCTGACTGGCAGAGAA | 1543 | | | | | | |
| DB | 8004 | CTTCTGAAGTCAAAGAAAAAATGGGAAAGTCTAAATGCTGTAACTAGCAATTTGGCAAAAGC | 8063 | | | | | | |
| QY | 1544 | AAATAGATGAGACCTTTGAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACC | 1603 | | | | | | |
| DB | 8064 | AAGTGGACAAGGCATTGGAGAAACTCAGAGACCTTGCAGGGAGCTATGGATGACCTGGACG | 8123 | | | | | | |
| QY | 1604 | TCAAGCTGCGCCAAGCTGAGTGTATCAAGGGATCCTGGCAGGCCCGTGGCGATCTCCTCA | 1663 | | | | | | |
| DB | 8124 | CTGACATGAAGGAGGCAGAGTCGTCGGAAATGGCTGGAAGCCCGTGGGAGACTTACTCA | 8183 | | | | | | |

```

RESULT 7
US-09-091-501B-5
; Sequence 5, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Rattus sp.
US-09-091-501B-5

```

| | Query Match | 4.0%; Score 79.4; DB 4; Length 200; | |
|----|--|-------------------------------------|--|
| | Best local Similarity 64.3%; Pred. NO. 7.9e-15; | | |
| | Matches 119; Conservative 0; Mismatches 66; Indels 0; Gaps 0; | | |
| QY | 465 CCTAAACGCCCAAGTACAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGT | 524 | |
| | | | |
| Db | 16 CCTGCAAAACCTGCTTGAAGAACATAAAAGTTTGCAAAAGTGACCTCGAAGCTGAGCAGGT | 75 | |
| | | | |
| QY | 525 CAGGGTCAATTCCTCACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGC | 584 | |
| | | | |
| Db | 76 GAAGGTGAATTCCTTAACCTCATATGGTGGTGGATTTGGATGAAAACAGTGGGGAGAGCGC | 135 | |
| | | | |
| QY | 585 AACTGCTGCTTTTGAAGAACAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAG | 644 | |
| | | | |
| Db | 136 CACAGCTGTTTTTGAAGATCAGTTACAGAAACTGGGTGAGCGCTGGACAGCTGTATGCCG | 195 | |
| | | | |
| QY | 645 ATGGA 649 | | |
| | | | |
| Db | 196 CTGGA 200 | | |
| | | | |

FILING DATE: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/935,313
 FILING DATE:
 APPLICATION NUMBER: EP 91 114 300.6
 FILING DATE: 26-AUG-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 30472/114 IMMU
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)836-9300
 TELEFAX: (703)683-4109
 TELEX: 899149
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7218 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 CLONE: PTZgpt-Fls
 US-08-232-463-14

| | Query Match | 3.8%; | Score 76.6; | DB 1; | Length 7218; |
|----|-----------------------|---|--------------------|------------|--------------|
| | Best Local Similarity | 6.7%; | Pred. No. 7.3e-13; | | |
| | Matches | Conservative | 236; | Mismatches | 155; |
| | | | | Indels | 0; |
| | | | | Gaps | 0; |
| QY | 113 | TTTCTAATGATGTGGAAGTGGTGAAGAGACCAGTTTCATACTCATGAGGGGTACATGATGG | 172 | | |
| DB | 1474 | TATCTATGCAAGTAGTTAAAGAGATAGAAGAAATTTGGTACRRRRRRRRRRRRRRRRRR | 1415 | | |
| QY | 173 | ATTGACAGCCCATCAGGGCCGGTTGGTAATAATTCACAATTTGGGAAGTAAGCTGATG | 232 | | |
| DB | 1414 | RR | 1355 | | |
| QY | 233 | GAACAGGAAAATTATCAGAAGATGAAGAACTGAAGTACAAGACAGATGAATCTCCTAA | 292 | | |
| DB | 1354 | RR | 1295 | | |
| QY | 293 | ATTCAAGATGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAG | 352 | | |
| DB | 1294 | RR | 1235 | | |
| QY | 353 | TTTTAATGGATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAG | 412 | | |
| DB | 1234 | RR | 1175 | | |
| QY | 413 | AAGAAAGAACAAAGGAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAAC | 472 | | |
| DB | 1174 | RR | 1115 | | |
| QY | 473 | GCCAAGTACAACAATAAGGTGCTTCAAGAAGATCTAGAAACAAGACAAGTCAGGTC | 531 | | |
| DB | 1114 | RR | 1056 | | |

RESULT 11

```

US-09-687-875A-13
; Sequence 13, Application US/09687875A
; Patent No. 6544786
; GENERAL INFORMATION:
; APPLICANT: Xiao, Xiao
; APPLICANT: Liu, Paul
; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPLICED PE
; FILE REFERENCE: 00792
; CURRENT APPLICATION NUMBER: US/09/687,875A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/158,868
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 238
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pXX-C2 5' junction
US-09-687-875A-13

```

| | Query Match | 3.2%; | Score 63.6; | DB 4; | Length 238; |
|----|-----------------------|--|---------------|-----------|-------------|
| | Best Local Similarity | 94.3%; | Pred. No. | 7.6e-10; | |
| | Matches 66; | Conservative 0; | Mismatches 4; | Indels 0; | Gaps 0; |
| QY | 1065 | TTCCGATGATGCAGTCTCTGTACAAAGACGTTTGATAACATGAACATTCAAGTGGAGTGA | 1124 | | |
| Dd | 169 | TTCGGACCGCAGTAGTCTGTACAAAGACGTTTGATAACATGAACATTCAAGTGGAGTGA | 228 | | |
| QY | 1125 | ACTTCGGAAA | 1134 | | |
| Dd | 229 | ACTTCGGAAA | 238 | | |

RESULT 12

US-09-198-452A-1/c
; Sequence 1, Application US/09198452A
; Patent No. 6559294

OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (300001)..(315000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (315001)..(330000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (330001)..(345000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (345001)..(360000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (360001)..(375000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (375001)..(390000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (390001)..(405000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (405001)..(420000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (420001)..(435000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (435001)..(450000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (450001)..(465000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (465001)..(480000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (480001)..(495000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (495001)..(510000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (510001)..(525000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (525001)..(540000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (540001)..(555000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (555001)..(570000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (570001)..(585000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (585001)..(600000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (600001)..(615000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (615001)..(630000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (630001)..(645000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (645001)..(660000)
OTHER INFORMATION: n=a or c or g or t

NAME/KEY: misc_feature
LOCATION: (660001)..(675000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (675001)..(690000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (690001)..(705000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (705001)..(720000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (720001)..(735000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (735001)..(750000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (750001)..(765000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (765001)..(780000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (780001)..(795000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (795001)..(810000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (810001)..(825000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (825001)..(840000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (840001)..(855000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (855001)..(870000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (870001)..(885000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (885001)..(900000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (900001)..(915000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature

Query Match 2.2%; Score 44; DB 4; Length 1230025;
Best Local Similarity 53.5%; Pred. No. 0.39;
Matches 92; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 242 AATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGAT 301
Db 656477 AATTAGAAGAAAGAGAAGAAATTGAGGATATCAAAAGACTCAGATACAAAAT 656418
QY 302 GGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTACATAGAGTTTAAATGG 361
Db 656417 GGGTTTCGATCACTCAAGCTGCTAAATTACATAACGCTAGGCAAGCAATTTATGTTG 656358
QY 362 ATCTCCAGATCAGAAACTGAAGAGTTGAATGACTGGCTAACAACAAAAACAGA 413
Db 656357 CAATTAGCAGAAAAAACTAAAGCTTCTAAAGAGACGCGCTGGGAAATAGA 656306

RESULT 13
US-09-107-532A-1186
; Sequence 1186, Application US/09107532A

RESULT 15
US-09-368-590-1
; Sequence 1, Application US/09368590
; Patent No. 6187563

GENERAL INFORMATION:
APPLICANT: Solimena, Michele
TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
TITLE OF INVENTION: AUTOANTIGENS OF AUTOIMMUNE DISEASES
FILE REFERENCE: 101918-200 (OCR-941)
CURRENT APPLICATION NUMBER: US/09/368,590
CURRENT FILING DATE: 1999-08-04
EARLIER APPLICATION NUMBER: 60/095,657
EARLIER FILING DATE: 1998-08-07
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 7812
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(6879)
NAME/KEY: unsure
LOCATION: (100)...(102)
NAME/KEY: unsure
LOCATION: (1021)...(1023)
NAME/KEY: unsure
LOCATION: (2266)...(2268)
JS-09-368-590-1

Query Match 2.1%; Score 42.8; DB 3; Length 7812;
Best Local Similarity 47.9%; Pred. No. 0.027;
Matches 156; Conservative 0; Mismatches 167; Indels 3; Gaps 1;
2y 1569 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGGCCAAGCTGAGGTGAT 1628
2b 3639 CAAGGAGTTGCACCAAGTGGCGCACGACCTGGACGAGCTGGCATGGGTTGAGGAGCG 3698
2y 1629 CAAGGGA---TCCTGGCAGCCGCTGGCGGATCTCTCATTTGACTCTCTCCAAGATCACCT 1685
2b 3699 GCTGCCACTGGCCATGCAGACAGAGCGAGGCAACGGTTTGACGGCGGTCCAGCAGCACAT 3758
2y 1686 CGAGAAAGTCAAGGCACCTTCAGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGT 1745
2b 3759 CAAAAGAACCAAGGCCCTGCGCGGGGAGATCCAGGCGATGGCGCGCTGGAGAGGT 3818
2y 1746 CAATGACCTTGCTGCCAGCTTACCACCTTTGGGCATTACAGCTCTCACCGTATAACCTCAG 1805
2b 3819 GCTGGAGCGCGGGCGGCTGCGTGGCAGCCCGGAGGCAGAGGAGTGGCGCG 3878
2y 1806 CACTCTGAAGACCTGAACACACAGATGGAAGCTTCTGAGGTGGCGGTGAGGACCGAGT 1865
2b 3879 GGGCCTGAGCAGCTGCAGAGCGCCTGGCCGGACTGCGGGAGGCTGCCGAGCGACGGCA 3938
2y 1866 CAGGAGCTGCATGAAGCCCAAGGG 1891
2b 3939 GCAGGTGCTGACCGCCCTTCCAGG 3964

Search completed: February 2, 2004, 06:15:49
Job time : 127.697 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

3M nucleic - nucleic search, using sw model

Run on: February 1, 2004, 10:13:25 ; Search time 475.292 Seconds
(without alignments)
11364.749 Million cell updates/sec

Title: US-09-845-416-6_COPY_1000_3000
Perfect score: 2001
Sequence: 1 ggcagttcattgatggagag.....tcaaccacgagactcaaac 2001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|--------------------|
| 1 | 2001 | 100.0 | 3999 | 24 | Human dystrophin m |
| 2 | 2001 | 100.0 | 4966 | 24 | Adeno-associated v |
| 3 | 2001 | 100.0 | 4990 | 24 | Adeno-associated v |
| 4 | 1709 | 85.4 | 3858 | 24 | Human dystrophin m |
| 5 | 1709 | 85.4 | 4825 | 24 | Adeno-associated v |
| 6 | 1709 | 85.4 | 4848 | 24 | Adeno-associated v |
| 7 | 1709 | 85.4 | 5060 | 24 | Adeno-associated v |
| 8 | 1662.8 | 83.1 | 4182 | 24 | Human dystrophin m |

| | | | | | | |
|----|--------|------|-------|----|----------|---------------------|
| 9 | 1662.8 | 83.1 | 5149 | 24 | AAD37255 | Adeno-associated v |
| 10 | 1283 | 64.1 | 5462 | 24 | ABK81999 | DNA encoding mini- |
| 11 | 1182.6 | 59.1 | 5952 | 22 | AAD06794 | Human dystrophin g |
| 12 | 1182.6 | 59.1 | 8689 | 24 | ABK82000 | DNA encoding mini- |
| 13 | 1182.6 | 59.1 | 11058 | 24 | AAD37229 | Human dystrophin p |
| 14 | 1182.6 | 59.1 | 11241 | 24 | ABK82005 | CDNA encoding huma |
| 15 | 1182.6 | 59.1 | 11443 | 24 | ABK82002 | DNA encoding mini- |
| 16 | 1182.6 | 59.1 | 12923 | 10 | AA903038 | Sequence of human |
| 17 | 1182.6 | 59.1 | 13957 | 24 | ABT10904 | Human breast cance |
| 18 | 1182.6 | 59.1 | 13957 | 24 | ABK69900 | Human dystrophin g |
| 19 | 1182.6 | 59.1 | 13957 | 24 | ABN95786 | Gene #2284 used to |
| 20 | 1182.6 | 59.1 | 13957 | 24 | ABK81959 | CDNA encoding huma |
| 21 | 1182.6 | 59.1 | 13977 | 24 | ABK70403 | Human bone remodel |
| 22 | 1182 | 59.1 | 2169 | 24 | AAD37232 | Human dystrophin r |
| 23 | 1182 | 59.1 | 3531 | 24 | AAD37238 | Human dystrophin m |
| 24 | 1182 | 59.1 | 4498 | 24 | AAD37258 | Adeno-associated v |
| 25 | 1180 | 59.0 | 5339 | 24 | ABK81998 | DNA encoding mini- |
| 26 | 1013 | 50.6 | 3510 | 24 | AAD37240 | Human dystrophin m |
| 27 | 1013 | 50.6 | 4476 | 24 | AAD37259 | Adeno-associated v |
| 28 | 1004 | 50.2 | 13815 | 19 | AAV18885 | Mus musculus dystro |
| 29 | 1004 | 50.2 | 13815 | 24 | ABK81960 | CDNA encoding mous |
| 30 | 1004 | 50.2 | 13815 | 24 | ABT99799 | Mouse ischaemic CO |
| 31 | 1004 | 50.2 | 19307 | 17 | AAT27558 | Shuttle vector pAd |
| 32 | 835 | 41.7 | 1821 | 24 | AAD37241 | Human dystrophin r |
| 33 | 727 | 36.3 | 5417 | 24 | ABK81997 | DNA encoding mini- |
| 34 | 678.4 | 33.9 | 4402 | 21 | AAZ48568 | A rod shortened dy |
| 35 | 678.4 | 33.9 | 4414 | 24 | AAD37260 | Adeno-associated v |
| 36 | 678 | 33.9 | 1991 | 24 | AAD37231 | Human dystrophin N |
| 37 | 677 | 33.8 | 1667 | 24 | AAD37235 | Human dystrophin N |
| 38 | 666.4 | 33.3 | 3446 | 24 | AAD37242 | Human dystrophin m |
| 39 | 618.8 | 30.9 | 4075 | 21 | AAZ48569 | A rod shortened dy |
| 40 | 614.2 | 30.7 | 4402 | 21 | AAZ48567 | A rod shortened dy |
| 41 | 537.8 | 26.9 | 3275 | 10 | AA97129 | Partial sequence o |
| 42 | 448 | 22.4 | 1434 | 24 | AAD37243 | Human dystrophin r |
| 43 | 408.2 | 20.4 | 11096 | 24 | ABK81962 | CDNA encoding mous |
| 44 | 397 | 19.8 | 6045 | 18 | AA74665 | Utrophin minigene. |
| 45 | 397 | 19.8 | 6059 | 22 | AA74673 | Nucleotide sequenc |

ALIGNMENTS

RESULT 1
AAD37234
ID AAD37234 standard; DNA; 3999 BP.

XX AAD37234;
XX 21-AUG-2002 (first entry)
XX Human dystrophin minigene delta3990.
XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 46-47; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3990 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1 and R2), 8059-10227 (rods R22, R23 and R24, hinge
CC H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids).
XX
SQ Sequence 3999 BP; 1223 A; 907 C; 933 G; 936 T; 0 other;
Query Match 100.0%; Score 2001; DB 24; Length 3999;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAA 60
Db 1000 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAA 1059
QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAAT 120
Db 1060 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAAT 1119
QY 121 GATGTGGAAGTGGTGAAGACCAAGTTTCATCTCATGAGGGGTACATGATGATTTGACA 180
Db 1120 GATGTGGAAGTGGTGAAGACCAAGTTTCATCTCATGAGGGGTACATGATGATTTGACA 1179
QY 181 GCCCATCAGGCCCGGTTGGTAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGA 240
Db 1180 GCCCATCAGGCCCGGTTGGTAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGA 1239
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 300
Db 1240 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 1299
QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATG 360
Db 1300 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATG 1359
QY 361 GATCTCCAGATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAAGAAAGA 420
Db 1360 GATCTCCAGATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAAGAAAGA 1419
QY 421 ACAAGGAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 480
Db 1420 ACAAGGAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 1479
QY 481 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGACAAGTCAGGGTCAATTCTCTC 540
Db 1480 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGACAAGTCAGGGTCAATTCTCTC 1539
QY 541 ACTCACATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
Db 1540 ACTCACATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1599
QY 601 GAACAACCTTAAGGTAATGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCG 660
Db 1600 GAACAACCTTAAGGTAATGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCG 1659
QY 661 TGGGTTCTTTTACAAGACCAAGCCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTCT 720
Db 1660 TGGGTTCTTTTACAAGACCAAGCCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTCT 1719

QY 721 CCTACTCAGACTGTACTCTGGTGACACAAACCTGTGGTTACTTAAGGAAACTGCCATCTCC 780
Db 1720 CCTACTCAGACTGTACTCTGGTGACACAAACCTGTGGTTACTTAAGGAAACTGCCATCTCC 1779
QY 781 AAAC TAGAAATGCCATCTTCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 840
Db 1780 AAAC TAGAAATGCCATCTTCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 1839
QY 841 TTCCCCCTGGACCTGGAAGAGTTTCTTTCCTGGCTTACAGAAGCTGAAACAACTGCCAAT 900
Db 1840 TTCCCCCTGGACCTGGAAGAGTTTCTTTCCTGGCTTACAGAAGCTGAAACAACTGCCAAT 1899
QY 901 GTCCTACAGGATGCTACCCGTAAGGAAAGCTCTCTAGAAAGACTCCAAAGGAGTAAAGAG 960
Db 1900 GTCCTACAGGATGCTACCCGTAAGGAAAGCTCTCTAGAAAGACTCCAAAGGAGTAAAGAG 1959
QY 961 CTGATGAAACAATGGCAAGACCTCCAAAGGTGAATTTGAAGCTCACACAGATGTTTATCAC 1020
Db 1960 CTGATGAAACAATGGCAAGACCTCCAAAGGTGAATTTGAAGCTCACACAGATGTTTATCAC 2019
QY 1021 AACCTGGATGAAACACGCCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTC 1080
Db 2020 AACCTGGATGAAACACGCCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTC 2079
QY 1081 CTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCT 1140
Db 2080 CTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCT 2139
QY 1141 CTCACATTAGGTCCCATTTGGAAAGCCAGTTCTGACAGTGGAAAGCGTCTGCACCTTTCT 1200
Db 2140 CTCACATTAGGTCCCATTTGGAAAGCCAGTTCTGACAGTGGAAAGCGTCTGCACCTTTCT 2199
QY 1201 CTGAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGAATTAAGCCGGCAGGCACCT 1260
Db 2200 CTGAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGAATTAAGCCGGCAGGCACCT 2259
QY 1261 ATTGAGGCGACTTTCCAGAGTTCAGAAGCAGAAACATGTACATAGGGCTTCAAGAGG 1320
Db 2260 ATTGAGGCGACTTTCCAGAGTTCAGAAGCAGAAACATGTACATAGGGCTTCAAGAGG 2319
QY 1321 GAATTGAAAACTAAAGAACCTGTAATCATGAGTACTTCTTGAGACTGTACGAATATTCTG 1380
Db 2320 GAATTGAAAACTAAAGAACCTGTAATCATGAGTACTTCTTGAGACTGTACGAATATTCTG 2379
QY 1381 ACAGAGCAGCCTTTGGAAGGACTAGAGAAAACCTCTACAGGAGCCAGAGAGCTGCCTCCT 1440
Db 2380 ACAGAGCAGCCTTTGGAAGGACTAGAGAAAACCTCTACAGGAGCCAGAGAGCTGCCTCCT 2439
QY 1441 GAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 1500
Db 2440 GAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 2499
QY 1501 GAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTT 1560
Db 2500 GAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTT 2559
QY 1561 GAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAGCT 1620
Db 2560 GAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAGCT 2619
QY 1621 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGAT 1680
Db 2620 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGAT 2679
QY 1681 CACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTGAGC 1740
Db 2680 CACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTGAGC 2739
QY 1741 CAGTCAATGACCTTGCTGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAAC 1800
Db 2740 CAGTCAATGACCTTGCTGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAAC 2799

Y 1801 CTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGCCGTCGAGGAC 1860
b 2800 CTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGCCGTCGAGGAC 2859
Y 1861 CGAGTCAGGCAGCTGCATGAAGCCACAGGACCTTGGTCCAGCATCTCAGCACCTTTCTT 1920
b 2860 CGAGTCAGGCAGCTGCATGAAGCCACAGGACCTTGGTCCAGCATCTCAGCACCTTTCTT 2919
Y 1921 TCCACGTCTGCCAGGTCCTGGGAGAGAGCCATCTGCCAAACAAAGTGCCCTACTAT 1980
b 2920 TCCACGTCTGCCAGGTCCTGGGAGAGAGCCATCTGCCAAACAAAGTGCCCTACTAT 2979
Y 1981 ATCAACCACGAGACTCAAAACA 2001
b 2980 ATCAACCACGAGACTCAAAACA 3000
RESULT 2
AD37256
D AAD37256 standard; DNA; 4966 BP.
X
C AAD37256;
T 21-AUG-2002 (first entry)
X Adeno-associated virus vector plasmid, AAV-MCK-delta3990.
E Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
W adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
W Becker muscular dystrophy; ds.
X Chimeric - Homo sapiens.
S Chimeric - Unidentified.
X WO200183695-A2.
X 08-NOV-2001.
X 27-APR-2001; 2001WO-US13677.
X 28-APR-2000; 2000US-200777P.
X (XIAO/) XIAO X.
X Xiao X;
X WPI; 2002-049342/06.
X New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
CX Example 1; Page 59-60; 7lpp; English.
CX The present invention relates to an isolated nucleotide sequence encoding
CX a dystrophin minigene. The minigene comprises N-terminal or modified
CX N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CX domains and cysteine-rich domains of dystrophin or utrophin genes. The
CX invention also relates to a recombinant adeno-associated virus (AAV)
CX comprising dystrophin minigene operably linked to an expression control
CX element. The dystrophin minigene in operable linkage with an expression
CX control element, in a recombinant adeno-associated virus or retrovirus is
CX useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CX dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CX vector plasmid construct containing human dystrophin minigenes, a muscle
CX creatine kinase (MCK) promoter and a small polyA signal sequence.
SQ Sequence 4966 BP; 1403 A; 1216 C; 1232 G; 1115 T; 0 other;
Query Match 100.0%; Score 2001; DB 24; Length 4966;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAA 60
Db 1757 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAA 1816
QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACACAGGAGAGATTTCTAAT 120
Db 1817 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACACAGGAGAGATTTCTAAT 1876
QY 121 GATGTGGAAGTGGTGAAGACCACTTTCATCTCATGAGGGGTACATGATGGATTGACA 180
Db 1877 GATGTGGAAGTGGTGAAGACCACTTTCATCTCATGAGGGGTACATGATGGATTGACA 1936
QY 181 GCCCATCAGGCCCGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240
Db 1937 GCCCATCAGGCCCGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 1996
QY 241 AAATTATCAGAAATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 300
Db 1997 AAATTATCAGAAATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 2056
QY 301 TGGGAATGCCCTCAGGCTAGCTAGCATGGAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAA 420
Db 2057 TGGGAATGCCCTCAGGCTAGCTAGCATGGAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAA 2176
QY 421 ACAAGGAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAACAGCCCAAGTA 480
Db 2177 ACAAGGAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAACAGCCCAAGTA 2236
QY 481 CAACAAACATAAGTGTCTTCAAGAAGATCTAGAACAAGAACCAAGTCAAGGTCAATTCTCTC 540
Db 2237 CAACAAACATAAGTGTCTTCAAGAAGATCTAGAACAAGAACCAAGTCAAGGTCAATTCTCTC 2296
QY 541 ACTCACATGGTGTGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGTGGAA 600
Db 2297 ACTCACATGGTGTGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGTGGAA 2356
QY 601 GAACAACTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCG 660
Db 2357 GAACAACTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCG 2416
QY 661 TGGGTTCTTTTACAAGACCAGCCTGACCTAGCTCTGACTGACCACTATTTGGAGCCTCT 720
Db 2417 TGGGTTCTTTTACAAGACCAGCCTGACCTAGCTCTGACTGACCACTATTTGGAGCCTCT 2476
QY 721 CCTACTCAGACTGTTTACTCTGGTGACACAACTGTGGTTACTAAGGAAACTGCCATCTCC 780
Db 2477 CCTACTCAGACTGTTTACTCTGGTGACACAACTGTGGTTACTAAGGAAACTGCCATCTCC 2536
QY 781 AAACCTAGAAATGCCATCTTCTTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 840
Db 2537 AAACCTAGAAATGCCATCTTCTTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 2596
QY 841 TTCCCTCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAAGCTGAAACAACTGCCAAT 900
Db 2597 TTCCCTCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAAGCTGAAACAACTGCCAAT 2656
QY 901 GTCTACAGGATGTACCCGTAAAGAAAGGCTCTTAGAAGACTCCAAGGGAGTAAAGAG 960
Db 2657 GTCTACAGGATGTACCCGTAAAGAAAGGCTCTTAGAAGACTCCAAGGGAGTAAAGAG 2716
QY 961 CTGATGAAACAATGGCAAGACCTCCAAGGTGAATTAAGACTCACACAGATGTTTATCAC 1020
Db 2717 CTGATGAAACAATGGCAAGACCTCCAAGGTGAATTAAGACTCACACAGATGTTTATCAC 2776
QY 1021 AACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCTCGAAGGTTCCGATGATGCAGTC 1080
Db 2777 AACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCTCGAAGGTTCCGATGATGCAGTC 2836


```
QY 1091 CTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAATAAGTCT 1140
Db      |||
QY 2837 CTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAATAAGTCT 2896
Db      |||
QY 1141 CTC AACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTCT 1200
Db      |||
QY 2897 CTC AACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTCT 2956
Db      |||
QY 1201 CTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGCGCAGGCACCT 1260
Db      |||
QY 2957 CTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGCGCAGGCACCT 3016
Db      |||
QY 1261 ATTGGAGGCGACTTTCCAGCAGTTTCAGAAGCAGACGATGTACATAGGGCCTTCAAGAGG 1320
Db      |||
QY 3017 ATTGGAGGCGACTTTCCAGCAGTTTCAGAAGCAGACGATGTACATAGGGCCTTCAAGAGG 3076
Db      |||
QY 1321 GAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTG 1380
Db      |||
QY 3077 GAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTG 3136
Db      |||
QY 1381 ACAGAGCAGCCCTTTGGAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCCTCCT 1440
Db      |||
QY 3137 ACAGAGCAGCCCTTTGGAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCCTCCT 3196
Db      |||
QY 1441 GAGGAGAGAGCCAGAAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 1500
Db      |||
QY 3197 GAGGAGAGAGCCAGAAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 3256
Db      |||
QY 1501 GAGTGGGAAAAAATTGAACCTTGGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 1560
Db      |||
QY 3257 GAGTGGGAAAAAATTGAACCTTGGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 3316
Db      |||
QY 1561 GAAAGACTCCAGGAACCTTCAAGAGGCCCACGGATGAGTGGACCTCAAGCTGCCCAAGCT 1620
Db      |||
QY 3317 GAAAGACTCCAGGAACCTTCAAGAGGCCCACGGATGAGTGGACCTCAAGCTGCCCAAGCT 3376
Db      |||
QY 1621 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGAT 1680
Db      |||
QY 3377 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGAT 3436
Db      |||
QY 1681 CACCTCGAGAAAGTCAAGGCACITCGAGGAGAAATTGGCGCTCTGAAAGAGAGAGCTGAGC 1740
Db      |||
QY 3437 CACCTCGAGAAAGTCAAGGCACITCGAGGAGAAATTGGCGCTCTGAAAGAGAGAGCTGAGC 3496
Db      |||
QY 1741 CACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAAC 1800
Db      |||
QY 3497 CACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAAC 3556
Db      |||
QY 1801 CTCAGCACITCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTGGCGCTCGAGGAC 1860
Db      |||
QY 3557 CTCAGCACITCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTGGCGCTCGAGGAC 3616
Db      |||
QY 1861 CGAGTCAGGACGCTGCATGAAGCCACAGGACTTTGGTCCAGCATCTCAGCACITTTCTT 1920
Db      |||
QY 3617 CGAGTCAGGACGCTGCATGAAGCCACAGGACTTTGGTCCAGCATCTCAGCACITTTCTT 3676
Db      |||
QY 1921 TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT 1980
Db      |||
QY 3677 TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT 3736
Db      |||
QY 1981 ATCAACCACGAGACTCAAAACA 2001
Db      |||
QY 3737 ATCAACCACGAGACTCAAAACA 3757
Db      |||
```

```
RESULT 3
AAD37262
ID AAD37262 standard; DNA; 4990 BP.
XX
AC AAD37262;
XX
DT 21-AUG-2002 (first entry)
XX
```

```
DE Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3990.
XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX Chimeric - Homo sapiens.
OS Chimeric - Cytomegalovirus.
OS Chimeric - Unidentified.
XX WO200183695-A2.
XX 08-NOV-2001.
XX 27-APR-2001; 2001WO-US13677.
XX 28-APR-2000; 2000US-200777P.
XX (XIAO/) XIAO X.
XX Xiao X;
XX WPI; 2002-049342/06.
XX
XX New dystrophin minigene for treating Duchenne or Becker muscular
XX dystrophy comprises an N-terminal domain or modified N-terminal domain,
XX rod repeats, H1 and H4 domains and a cysteine rich domain of a
XX dystrophin gene -
XX
XX Example 1; Page 67-68; 71pp; English.
XX
XX The present invention relates to an isolated nucleotide sequence encoding
XX a dystrophin minigene. The minigene comprises N-terminal or modified
XX N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
XX domains and cysteine-rich domains of dystrophin or utrophin genes. The
XX invention also relates to a recombinant adeno-associated virus (AAV)
XX comprising dystrophin minigene operably linked to an expression control
XX element. The dystrophin minigene in operable linkage with an expression
XX control element, in a recombinant adeno-associated virus or retrovirus is
XX useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
XX dystrophy (BMD) in a mammalian subject. The present sequence is AAV
XX vector plasmid construct containing human dystrophin minigenes, a
XX cytomegalovirus (CMV) promoter and a small polyA signal sequence.
XX
XX Sequence 4990 BP; 1439 A; 1185 C; 1208 G; 1158 T; 0 other;
XX
XX Query Match 100.0%; Score 2001; DB 24; Length 4990;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAA 60
Db      |||
QY 1781 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAA 1840
Db      |||
QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTCGAAGCACAAAGGAGAGATTCTAAT 120
Db      |||
QY 1841 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTCGAAGCACAAAGGAGAGATTCTAAT 1900
Db      |||
QY 121 GATGTGGAAGTGGTGAAGACCCAGTTTCATCTCATGAGGGGTACATGATGGATTGACA 180
Db      |||
QY 1901 GATGTGGAAGTGGTGAAGACCCAGTTTCATCTCATGAGGGGTACATGATGGATTGACA 1960
Db      |||
QY 181 GCCCATCAGGCCCGGTTGGTAATATTTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGA 240
Db      |||
QY 1961 GCCCATCAGGCCCGGTTGGTAATATTTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGA 2020
Db      |||
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGATCTCTCTAAATTCAAGA 300
Db      |||
QY 2021 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGATCTCTCTAAATTCAAGA 2080
Db      |||
QY 301 TGGGAATGCCTCAGGGTAGCTAGCTAGTGAAGAAACAAAGCAATTACATAGAGTTTAAATG 360
Db      |||
QY 2081 TGGGAATGCCTCAGGGTAGCTAGCTAGTGAAGAAACAAAGCAATTACATAGAGTTTAAATG 2140
Db      |||
```

361 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGA 420
1b
2141 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGA 2200
421 ACAAGGAAAAATGGAGAAAGAGCCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTA 480
1b
2201 ACAAGGAAAAATGGAGAAAGAGCCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTA 2260
481 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTC 540
1b
2261 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTC 2320
541 ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 600
1b
2321 ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 2380
601 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGC 660
1b
2381 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGC 2440
661 TGGGTTCTTTTACAAGACAGCCCTGACCTAGCTCTCTGGACTGACCACTATTGGAGCCTCT 720
1b
2441 TGGGTTCTTTTACAAGACAGCCCTGACCTAGCTCTCTGGACTGACCACTATTGGAGCCTCT 2500
721 CCTACTCAGACTGTTACTCTGGTGACACAACCTGTGGTTACTTAAGGAACTGCCATCTCC 780
1b
2501 CCTACTCAGACTGTTACTCTGGTGACACAACCTGTGGTTACTTAAGGAACTGCCATCTCC 2560
781 AAAGTAGAAATGCCATCTTCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 840
1b
2561 AAAGTAGAAATGCCATCTTCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 2620
841 TTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAAGCTGAAACAACTGCCAAT 900
1b
2621 TTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAAGCTGAAACAACTGCCAAT 2680
901 GTCTACAGGATGCTACCCGTAAAGAAAGGCTCCTAGAAAGCTCCAAGGGAGTAAAGAG 960
1b
2681 GTCTACAGGATGCTACCCGTAAAGAAAGGCTCCTAGAAAGCTCCAAGGGAGTAAAGAG 2740
961 CTGATGAACAAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTATCAC 1020
1b
2741 CTGATGAACAAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTATCAC 2800
1021 AACCTGGATGAAAAACAGCCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGAGTC 1080
1b
2801 AACCTGGATGAAAAACAGCCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGAGTC 2860
1081 CTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAATTCGGAAAAAGTCT 1140
1b
2861 CTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAATTCGGAAAAAGTCT 2920
1141 CTCAACATTAGTCCCATTTTGGAAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCT 1200
1b
2921 CTCAACATTAGTCCCATTTTGGAAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCT 2980
1201 CTGCAGGAACCTTCTGGCTACAGCTGAAAGATGATGAATTAAGCCCGCAGGCACCT 1260
1b
2981 CTGCAGGAACCTTCTGGCTACAGCTGAAAGATGATGAATTAAGCCCGCAGGCACCT 3040
1261 ATTGGAGGCGACTTCCAGCAGTTTCAGAACAGAACCGATGTACATAGGGCCTTCAAGAGG 1320
1b
3041 ATTGGAGGCGACTTCCAGCAGTTTCAGAACAGAACCGATGTACATAGGGCCTTCAAGAGG 3100
1321 GAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 1380
1b
3101 GAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 3160
1381 ACAGAGCAGCCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCCGCAGAGAGCTGCCTCT 1440
1b
3161 ACAGAGCAGCCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCCGCAGAGAGCTGCCTCT 3220

1441 GAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 1500
1b
3221 GAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 3280
1501 GAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 1560
1b
3281 GAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 3340
1561 GAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA CCTCAAGCTGCGCCAAGCT 1620
1b
3341 GAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA CCTCAAGCTGCGCCAAGCT 3400
1621 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGAT 1680
1b
3401 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGAT 3460
1681 CACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTTGAGC 1740
1b
3461 CACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTTGAGC 3520
1741 CACGTCAATGACCTTGTCTCGCCAGCTTACCCTTTGGGCATTTCAGCTCTCACCGTATAAC 1800
1b
3521 CACGTCAATGACCTTGTCTCGCCAGCTTACCCTTTGGGCATTTCAGCTCTCACCGTATAAC 3580
1801 CTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAAGTGGCGCTCGAGGAC 1860
1b
3581 CTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAAGTGGCGCTCGAGGAC 3640
1861 CGAGTCAGGCAGCTGCATGAAGCCCAAGGACCTTGGTCCAGCATCTCAGCACTTTCTT 1920
1b
3641 CGAGTCAGGCAGCTGCATGAAGCCCAAGGACCTTGGTCCAGCATCTCAGCACTTTCTT 3700
1921 TCCACGTCTGTCCAGGTCCTGGGAGAGAGCCCATCTCGCCAAAACAAAGTGCCTACTAT 1980
1b
3701 TCCACGTCTGTCCAGGTCCTGGGAGAGAGCCCATCTCGCCAAAACAAAGTGCCTACTAT 3760
1981 ATCAACCACGAGACTCAAACA 2001
1b
3761 ATCAACCACGAGACTCAAACA 3781

RESULT 4

AAD37237
ID AAD37237 standard; DNA; 3858 BP.

XX
AC AAD37237;

XX
DT 21-AUG-2002 (first entry)

XX
DE Human dystrophin minigene delta3849.

XX
DE Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

KW Becker muscular dystrophy; ds.

XX
OS .Homo sapiens.

XX
PN WO200183695-A2.

XX
PD 08-NOV-2001.

XX
PF 27-APR-2001; 2001WO-US13677.

XX
PR 28-APR-2000; 2000US-200777P.

XX
PA (XIAO/) XIAO X.

XX
PI Xiao X;

XX
DR WPI; 2002-049342/06.

XX
PT New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,

rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene -

Example 1; Page 48-49; 71pp; English.

The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is human dystrophin minigene delta3849 containing nucleotides 1-1668 (N-terminus, hinge H1 and rods R1, R2), 8059-10227 (rods R22, R23 and R24, hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids).

Sequence 3858 BP; 1189 A; 866 C; 905 G; 898 T; 0 other;

Query Match 85.4%; Score 1709; DB 24; Length 3858; Best Local Similarity 93.0%; Pred. No. 0; Matches 1860; Conservative 0; Mismatches 0; Indels 141; Gaps 1;

| | | | |
|----|------|--|------|
| QY | 1 | GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAA | 60 |
| Db | 1000 | GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAA | 1059 |
| QY | 61 | GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAAT | 120 |
| Db | 1060 | GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAAT | 1119 |
| QY | 121 | GATGTGAAGTGGTGAAAGACACAGTTTCATCTCATGAGGGGTACATGATGATTTGACA | 180 |
| Db | 1120 | GATGTGAAGTGGTGAAAGACACAGTTTCATCTCATGAGGGGTACATGATGATTTGACA | 1179 |
| QY | 181 | GCCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA | 240 |
| Db | 1180 | GCCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA | 1239 |
| QY | 241 | AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTTAATTCAAGA | 300 |
| Db | 1240 | AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTTAATTCAAGA | 1299 |
| QY | 301 | TGGGAATGCCTCAGGGTAGCTAGCATGGGAAAAACAAAGCAATTTACATAGAGTTTAATG | 360 |
| Db | 1300 | TGGGAATGCCTCAGGGTAGCTAGCATGGGAAAAACAAAGCAATTTACATAGAGTTTAATG | 1359 |
| QY | 361 | GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAAGA | 420 |
| Db | 1360 | GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAAGA | 1419 |
| QY | 421 | ACAAGGAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA | 480 |
| Db | 1420 | ACAAGGAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA | 1479 |
| QY | 481 | CAACAACATAAGTGCTTCAAGAAGATCTAGAACAAGAACAAAGTCAAGGTCAATTCTCTC | 540 |
| Db | 1480 | CAACAACATAAGTGCTTCAAGAAGATCTAGAACAAGAACAAAGTCAAGGTCAATTCTCTC | 1539 |
| QY | 541 | ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA | 600 |
| Db | 1540 | ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA | 1599 |
| QY | 601 | GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGC | 660 |
| Db | 1600 | GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGC | 1659 |
| QY | 661 | TGGGTTCTTTTACAAGACACCGCTGACCTCTGGACTGCCACTATTGGAGCCTCT | 720 |
| Db | 1660 | TGGGTTCTTTTACAAGAC----- | 1677 |

| | | | |
|----|------|--|------|
| QY | 721 | CCTACTCAGACTGTTACTCTGGTGACACAACCTGTGGTTACTAAGAAACTGCCATCTCC | 780 |
| Db | 1678 | ----- | 1677 |
| QY | 781 | AAACTAGAAATGCCATCTTCTTGATGTGGAGGTACCTACTCATAGATTACTGCAACAG | 840 |
| Db | 1678 | -----ACTCATAGATTACTGCAACAG | 1698 |
| QY | 841 | TTCCCCCTGGACCTGGAAGAAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACTGCCAAT | 900 |
| Db | 1699 | TTCCCCCTGGACCTGGAAGAAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACTGCCAAT | 1758 |
| QY | 901 | GTCTTACAGGATGTACCCGTAAAGAAAGGCTCCTAGAAAGACTCCAAGGGAGTAAAGAG | 960 |
| Db | 1759 | GTCTTACAGGATGTACCCGTAAAGAAAGGCTCCTAGAAAGACTCCAAGGGAGTAAAGAG | 1818 |
| QY | 961 | CTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCAC | 1020 |
| Db | 1819 | CTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCAC | 1878 |
| QY | 1021 | AACCTGGATGAAACACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTC | 1080 |
| Db | 1879 | AACCTGGATGAAACACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTC | 1938 |
| QY | 1081 | CTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAGTCT | 1140 |
| Db | 1939 | CTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAGTCT | 1998 |
| QY | 1141 | CTCAACATTAGTCCCATTTTGGAAAGCCAGTTCTGACACAGTGGAGCGTCTGCACCTTTCT | 1200 |
| Db | 1999 | CTCAACATTAGTCCCATTTTGGAAAGCCAGTTCTGACACAGTGGAGCGTCTGCACCTTTCT | 2058 |
| QY | 1201 | CTGCAGGAACCTTCTGGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT | 1260 |
| Db | 2059 | CTGCAGGAACCTTCTGGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT | 2118 |
| QY | 1261 | ATTGGAGCGGACTTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGGGCTTCAAGAGG | 1320 |
| Db | 2119 | ATTGGAGCGGACTTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGGGCTTCAAGAGG | 2178 |
| QY | 1321 | GAATTGAAAACTAAAGAACCTGTATCATGACTACTCTTGAGACTGTACGAATATTTCTG | 1380 |
| Db | 2179 | GAATTGAAAACTAAAGAACCTGTATCATGACTACTCTTGAGACTGTACGAATATTTCTG | 2238 |
| QY | 1381 | ACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACAGAGAGCCAGAGAGCTGCCTCCT | 1440 |
| Db | 2239 | ACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACAGAGAGCCAGAGAGCTGCCTCCT | 2298 |
| QY | 1441 | GAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT | 1500 |
| Db | 2299 | GAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT | 2358 |
| QY | 1501 | GAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT | 1560 |
| Db | 2359 | GAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT | 2418 |
| QY | 1561 | GAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCT | 1620 |
| Db | 2419 | GAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCT | 2478 |
| QY | 1621 | GAGGTGATCAAGGGATCCTGGCAGCCGCTGGCGGATCTCTCATTTGACTCTCTCCAAGAT | 1680 |
| Db | 2479 | GAGGTGATCAAGGGATCCTGGCAGCCGCTGGCGGATCTCTCATTTGACTCTCTCCAAGAT | 2538 |
| QY | 1681 | CACCTCGAGAAAGTCAAGGCACCTCGAGGAGAAATTGCGCTCTGAAAGAGAACGTGAGC | 1740 |
| Db | 2539 | CACCTCGAGAAAGTCAAGGCACCTCGAGGAGAAATTGCGCTCTGAAAGAGAACGTGAGC | 2598 |
| QY | 1741 | CACGTCAATGACCTTGCTCGCCAGTTTACCACCTTTGGGATTCAGCTCTCACCGTATAAC | 1800 |
| Db | 2599 | CACGTCAATGACCTTGCTCGCCAGTTTACCACCTTTGGGATTCAGCTCTCACCGTATAAC | 2658 |
| QY | 1801 | CTCAGCACTCTGGAAGACCTTGAAACACACAGATGGAAGCTTTCTGAGGTGGCCGTGAGGAC | 1860 |

Db 2696 CTGTTACAAAGACGTTTGGATACATGAAGTCAAGTGGAGTGAAGTTCGGAAAAAGTCT 2755
QY 1141 CTCAACATTAGTCCCATTTGGAAGCCAGTCTGACCAGTGAAGCGTCTGCACCTTTCT 1200
Db 2756 CTCAACATTAGTCCCATTTGGAAGCCAGTCTGACCAGTGAAGCGTCTGCACCTTTCT 2815
QY 1201 CTGCAGGAACCTTCTGGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCT 1260
Db 2816 CTGCAGGAACCTTCTGGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCT 2875
QY 1261 ATTGAGGCGACTTTCCAGCAGTTTCAGAACGATGTACATAGGCGCTTCAAGAGG 1320
Db 2876 ATTGAGGCGACTTTCCAGCAGTTTCAGAACGATGTACATAGGCGCTTCAAGAGG 2935
QY 1321 GAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTCTG 1380
Db 2936 GAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTCTG 2995
QY 1381 ACAGAGCAGCCTTTTGAAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCCTCCT 1440
Db 2996 ACAGAGCAGCCTTTTGAAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCCTCCT 3055
QY 1441 GAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATACT 1500
Db 3056 GAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATACT 3115
QY 1501 GAGTGGAAAAATTGAACCTGCACTCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 1560
Db 3116 GAGTGGAAAAATTGAACCTGCACTCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 3175
QY 1561 GAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCGCCAAGCT 1620
Db 3176 GAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCGCCAAGCT 3235
QY 1621 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGAT 1680
Db 3236 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGAT 3295
QY 1681 CACCTCGAGAAAGTCAAGGCACCTTCAGAGGAGAAATTGGCGCTCTGAAAGAGAACGTGAGC 1740
Db 3296 CACCTCGAGAAAGTCAAGGCACCTTCAGAGGAGAAATTGGCGCTCTGAAAGAGAACGTGAGC 3355
QY 1741 CAGCTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAAC 1800
Db 3356 CAGCTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAAC 3415
QY 1801 CTCAGACTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTCAGGTTGGCCGTCGAGGAC 1860
Db 3416 CTCAGACTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTCAGGTTGGCCGTCGAGGAC 3475
QY 1861 CGAGTCAGGAGCTGCATGAAGCCACACGGGACTTTGGTCCAGCATCTCAGCACTTTCTT 1920
Db 3476 CGAGTCAGGAGCTGCATGAAGCCACACGGGACTTTGGTCCAGCATCTCAGCACTTTCTT 3535
QY 1921 TCCACGTCTGTCCAGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTCCCTACTAT 1980
Db 3536 TCCACGTCTGTCCAGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTCCCTACTAT 3595
QY 1981 ATCAACCACGAGACTCAAAACA 2001
Db 3596 ATCAACCACGAGACTCAAAACA 3616

RESULT 6
AAD37263
ID AAD37263 standard; DNA; 4848 BP.
XX
AC AAD37263;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3849.

XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Cytomegalovirus.
OS Chimeric - Unidentified.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 68-70; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC cytomegalovirus (CMV) promoter and a small polyA signal sequence.
XX
SQ Sequence 4848 BP; 1405 A; 1144 C; 1180 G; 1119 T; 0 other;

Query Match 85.4%; Score 1709; DB 24; Length 4848;
Best Local Similarity 93.0%; Pred. No. 0;
Matches 1860; Conservative 0; Mismatches 0; Indels 141; Gaps 1;

QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA 60
Db 1780 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA 1839
QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAAT 120
Db 1840 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAAT 1899
QY 121 GATGTGGAAGTGGTGAAGACCAAGTTCATCTACTCATGAGGGGTACATGATGGATTGACA 180
Db 1900 GATGTGGAAGTGGTGAAGACCAAGTTCATCTACTCATGAGGGGTACATGATGGATTGACA 1959
QY 181 GCCCATCAGGCGCGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240
Db 1960 GCCCATCAGGCGCGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 2019
QY 241 AAATTATCAGAGATGAAGAAACCTGAAGTACAAAGAGCAGATGAATCTCCTAAATTCAAGA 300
Db 2020 AAATTATCAGAGATGAAGAAACCTGAAGTACAAAGAGCAGATGAATCTCCTAAATTCAAGA 2079
QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATG 360
Db 2080 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATG 2139

2Y 361 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAA 420
2b 2140 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAA 2199
2Y 421 ACAAGGAAAAATGGAGGAGAGCCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAAGTA 480
2b 2200 ACAAGGAAAAATGGAGGAGAGCCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAAGTA 2259
2Y 481 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACACAGAAACAAGTCAGGGTCAATTCTCTC 540
2b 2260 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACACAGAAACAAGTCAGGGTCAATTCTCTC 2319
2Y 541 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGTGGAA 600
2b 2320 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGTGGAA 2379
2Y 601 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGC 660
2b 2380 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGC 2439
2Y 661 TGGGTCTTTTACAAGACCCGCTGACCTAGCTCCTGSACTGACCATTATTGGAGCCTCT 720
2b 2440 TGGGTCTTTTACAAGAC----- 2457
2Y 721 CCTACTCAGACTGTTACTCTGGTGACACAACTGTGGTTACTAAGGAAACTGSCATCTCC 780
2b 2458 ----- 2457
2Y 781 AAAC TAGAAATGCCATCTTCCCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 840
2b 2458 -----ACTCATAGATTACTGCAACAG 2478
2Y 841 TTCCCTCCCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGAAAGCTGAAACAACTGCCAAT 900
2b 2479 TTCCCTCCCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGAAAGCTGAAACAACTGCCAAT 2538
2Y 901 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAAGCTCCAAGGGAGTAAAGAG 960
2b 2539 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAAGCTCCAAGGGAGTAAAGAG 2598
2Y 961 CTGATGAAACAAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCAC 1020
2b 2599 CTGATGAAACAAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCAC 2658
2Y 1021 AACCTGGATGAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTC 1080
2b 2659 AACCTGGATGAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTC 2718
2Y 1081 CTGTTACAAAGACGTTTGGATAACATGAATCTCAAGTGGAGTGAACTTCGGAAAAAGTCT 1140
2b 2719 CTGTTACAAAGACGTTTGGATAACATGAATCTCAAGTGGAGTGAACTTCGGAAAAAGTCT 2778
2Y 1141 CTCACAATTAGTCCCATTTTGGAAAGCCAGTTCTGACCAGTGGAGCGTCTGCACCTTTCT 1200
2b 2779 CTCACAATTAGTCCCATTTTGGAAAGCCAGTTCTGACCAGTGGAGCGTCTGCACCTTTCT 2838
2Y 1201 CTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT 1260
2b 2839 CTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCT 2898
2Y 1261 ATTGGAGGCGACTTTCCAGCAGTTCCAGAACAGAACGATGTACATAGGGCCTTCAAGAGG 1320
2b 2899 ATTGGAGGCGACTTTCCAGCAGTTCCAGAACAGAACGATGTACATAGGGCCTTCAAGAGG 2958
2Y 1321 GAATTGAAAACTAAAGAACCTGTAAATCATAGTACTCTTTGAGACTGTACGAATATTTCTG 1380
2b 2959 GAATTGAAAACTAAAGAACCTGTAAATCATAGTACTCTTTGAGACTGTACGAATATTTCTG 3018
2Y 1381 ACAGAGCAGCCTTTGGAGGACTAGAGAACTCTACAGGAGCCACAGAGAGCTGCCTCCT 1440
2b 3019 ACAGAGCAGCCTTTGGAGGACTAGAGAACTCTACAGGAGCCACAGAGAGCTGCCTCCT 3078
2Y 1441 GAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 1500

Db 3079 GAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACT 3138
QY 1501 GAGTGGGAAAAAATTGAACCTGCACCTCCGCTGACTCCGCTGAGTGGCAGAGAAAAATAGATGAGACCCCTT 1560
Db 3139 GAGTGGGAAAAAATTGAACCTGCACCTCCGCTGACTCCGCTGAGTGGCAGAGAAAAATAGATGAGACCCCTT 3198
QY 1561 GAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCT 1620
Db 3199 GAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCT 3258
QY 1621 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCTCATTTGAGTCTCTCCAAAGAT 1680
Db 3259 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCTCATTTGAGTCTCTCCAAAGAT 3318
QY 1681 CACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAAGAGAACGTGAGC 1740
Db 3319 CACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAAGAGAACGTGAGC 3378
QY 1741 CACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCACTTCAGCTCTCACCGTATAAC 1800
Db 3379 CACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCACTTCAGCTCTCACCGTATAAC 3438
QY 1801 CTCAGCACTCTGGAAGACCTTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGAC 1860
Db 3439 CTCAGCACTCTGGAAGACCTTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGAC 3498
QY 1861 CGAGTCAGGCAGCTGCATGAAGCCACAGGGACCTTTGGTCCAGCATCTCAGCACCTTTCTT 1920
Db 3499 CGAGTCAGGCAGCTGCATGAAGCCACAGGGACCTTTGGTCCAGCATCTCAGCACCTTTCTT 3558
QY 1921 TCCACGTCTGTCCAGGTCCTCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT 1980
Db 3559 TCCACGTCTGTCCAGGTCCTCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT 3618
QY 1981 ATCAACCACGAGACTCAAACA 2001
Db 3619 ATCAACCACGAGACTCAAACA 3639

RESULT 7

AAD37264
ID AAD37264 standard; DNA; 5060 BP.

XX AAD37264;

XX 21-AUG-2002 (first entry)

XX Adeno-associated virus (AAV) vector plasmid, AAV-B-CMV-3849.

DE Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.

XX Chimeric - Homo sapiens.
OS Chimeric - Cytomegalovirus.
OS Chimeric - Unidentified.

XX WO200183695-A2.

XX 08-NOV-2001.

XX 27-APR-2001; 2001WO-US13677.

XX 28-APR-2000; 2000US-200777P.

XX (XIAO/) XIAO X.

XX Xiao X;

XX WPI; 2002-049342/06.

XX New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 70-71; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC muscle creatine kinase (MCK) enhancer, a cytomegalovirus (CMV) promoter
CC and a small polyA signal sequence.
XX
SQ Sequence 5060 BP; 1449 A; 1217 C; 1234 G; 1160 T; 0 other;

Query Match 85.4%; Score 1709; DB 24; Length 5060;
Best Local Similarity 93.0%; Pred. No. 0;
Matches 1860; Conservative 0; Mismatches 0; Indels 141; Gaps 1;

QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAA 60
DB 1992 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAA 2051

QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTAAT 120
DB 2052 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTAAT 2111

QY 121 GATGTGGAAGTGGTGAAAGACCAAGTTCATCTCATGAGGGGTACATGATGGATTGACA 180
DB 2112 GATGTGGAAGTGGTGAAAGACCAAGTTCATCTACTCATGAGGGGTACATGATGGATTGACA 2171

QY 181 GCCCATCAGGCCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGA 240
DB 2172 GCCCATCAGGCCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGA 2231

QY 241 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 300
DB 2232 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 2291

QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAACAAAGCAATTTACATAGAGTTTAAATG 360
DB 2292 TGGGAATGCCTCAGGGTAGCTAGCATGGAAACAAAGCAATTTACATAGAGTTTAAATG 2351

QY 361 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAACACAGAAGAAAGA 420
DB 2352 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAACACAGAAGAAAGA 2411

QY 421 ACAAGGAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTGAAGACTTAACAAACGCCAAGTA 480
DB 2412 ACAAGGAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTGAAGACTTAACAAACGCCAAGTA 2471

QY 481 CAACAACATAAGTGCTTCAAGAAGATCTAGAACAAGAACAAAGTCAAGGTCAATTCTCTC 540
DB 2472 CAACAACATAAGTGCTTCAAGAAGATCTAGAACAAGAACAAAGTCAAGGTCAATTCTCTC 2531

QY 541 ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 600
DB 2532 ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 2591

QY 601 GAACAACCTTAAGGTATTGGAGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGC 660
DB 2592 GAACAACCTTAAGGTATTGGAGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGC 2651

QY 661 TGGGTCTTTTACAAGACCCAGCTGACCTAGCTCCTGGACTGACCCTATTGGAGCCTCT 720
DB 2652 TGGGTCTTTTACAAGAC----- 2669

QY 721 CCTACTCAGACTGTTACTCTGGTGACACAACCTGTGTTACTAAGGAAACTGCCATCTCC 780
DB 2670 ----- 2669

QY 781 AAACCTAGAAATGCCATCTTCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 840
DB 2670 -----ACTCATAGATTACTGCAACAG 2690

QY 841 TTCCCCCTGGACCTGGAAAAGTTTCTTCTGCTGGCTTACAGAAGCTGAAACAACTGCCAAT 900
DB 2691 TTCCCCCTGGACCTGGAAAAGTTTCTTCTGCTGGCTTACAGAAGCTGAAACAACTGCCAAT 2750

QY 901 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAAGACTCCAAAGGAGTAAAAGAG 960
DB 2751 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAAGACTCCAAAGGAGTAAAAGAG 2810

QY 961 CTGATGAAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCAC 1020
DB 2811 CTGATGAAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCAC 2870

QY 1021 AACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTC 1080
DB 2871 AACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTC 2930

QY 1081 CTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAGTCT 1140
DB 2931 CTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAGTCT 2990

QY 1141 CTCAACATTTAGGTCCCAATTTGGAAGCCAGTTCTGACCAAGTGGAAAGCGTCTGCACCTTTCT 1200
DB 2991 CTCAACATTTAGGTCCCAATTTGGAAGCCAGTTCTGACCAAGTGGAAAGCGTCTGCACCTTTCT 3050

QY 1201 CTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCCGCAGGCACCT 1260
DB 3051 CTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCCGCAGGCACCT 3110

QY 1261 ATTGGAGCGGACTTTCCAGCAGTTTCCAGAGCAGAACAGATGATGATAGGGCTTCAAGAGG 1320
DB 3111 ATTGGAGCGGACTTTCCAGCAGTTTCCAGAGCAGAACAGATGATGATAGGGCTTCAAGAGG 3170

QY 1321 GAATTGAAAAACTAAAGAACCTGTAATCATGATGATGATGATGATGATGATGATGATGATG 1380
DB 3171 GAATTGAAAAACTAAAGAACCTGTAATCATGATGATGATGATGATGATGATGATGATG 3230

QY 1381 ACAGAGCAGCCTTTTGAAGGACTAGAGAAACTCTACCAAGGAGCCAGAGAGCTGCCTCCT 1440
DB 3231 ACAGAGCAGCCTTTTGAAGGACTAGAGAAACTCTACCAAGGAGCCAGAGAGCTGCCTCCT 3290

QY 1441 GAGGAGAGAGCCAGAAATGTCTACCTCGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 1500
DB 3291 GAGGAGAGAGCCAGAAATGTCTACCTCGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 3350

QY 1501 GAGTGGGAAAAAATTGAACCTGCCTCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 1560
DB 3351 GAGTGGGAAAAAATTGAACCTGCCTCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTT 3410

QY 1561 GAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGGCCAAGCT 1620
DB 3411 GAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGGCCAAGCT 3470

QY 1621 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGAT 1680
DB 3471 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGAT 3530

QY 1681 CACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTGAGC 1740
DB 3531 CACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTGAGC 3590

QY 1741 CACGTCAATGACCTTGTCTGCCAGCTTACCCTTTGGGCTTTCAGCTCTCACCGTATAAC 1800
DB 3591 CACGTCAATGACCTTGTCTGCCAGCTTACCCTTTGGGCTTTCAGCTCTCACCGTATAAC 3650

QY 898 AATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAGACTCCAAGGGAGTAAA 957
Db
QY 958 GAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAATGAAGCTCACACAGATGTTTAT 1017
Db 2140 GAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAATGAAGCTCACACAGATGTTTAT 2199
QY 1018 CACAACCTGGATGAAAACAGCCCAAAAAATCCTGAGATCCCTGGAAAGTTCCGATGATGCA 1077
Db 2200 CACAACCTGGATGAAAACAGCCCAAAAAATCCTGAGATCCCTGGAAAGTTCCGATGATGCA 2259
QY 1078 GTCCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAG 1137
Db 2260 GTCCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAG 2319
QY 1138 TCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTT 1197
Db 2320 TCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTT 2379
QY 1198 TCTCTGAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCA 1257
Db 2380 TCTCTGAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCA 2439
QY 1258 CCTATTGGAGGCGACTTTCAGCAGTTTCCAGAGCAGAACGATGTACATAGGCGCTTCAAG 1317
Db 2440 CCTATTGGAGGCGACTTTCAGCAGTTTCCAGAGCAGAACGATGTACATAGGCGCTTCAAG 2499
QY 1318 AGGGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTTGAGACTGTACGAATATTT 1377
Db 2500 AGGGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTTGAGACTGTACGAATATTT 2559
QY 1378 CTGACAGAGCAGCCCTTTTGGAGGACTAGAGAAACTCTACAGAGAGCCAGAGAGCTGCCT 1437
Db 2560 CTGACAGAGCAGCCCTTTTGGAGGACTAGAGAAACTCTACAGAGAGCCAGAGAGCTGCCT 2619
QY 1438 CCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAAT 1497
Db 2620 CCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAAT 2679
QY 1498 ACTGAGTGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACC 1557
Db 2680 ACTGAGTGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACC 2739
QY 1558 CTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAA 1617
Db 2740 CTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAA 2799
QY 1618 GCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCTCATTTGACTCTCTCCAA 1677
Db 2800 GCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCTCATTTGACTCTCTCCAA 2859
QY 1678 GATCACCTCGAGAAAGTCAAGGACCTTCGAGGAGAAATTGCGGCTCTGAAAGAGAACGTG 1737
Db 2860 GATCACCTCGAGAAAGTCAAGGACCTTCGAGGAGAAATTGCGGCTCTGAAAGAGAACGTG 2919
QY 1738 AGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTAT 1797
Db 2920 AGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTAT 2979
QY 1798 AACCTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGAGGTGGCCGTCGAG 1857
Db 2980 AACCTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGAGGTGGCCGTCGAG 3039
QY 1858 GACCGAGTCAGGAGTGCATGAAGCCCAAGGGACTTTGGTCCAGCATCTCAGCACTT 1917
Db 3040 GACCGAGTCAGGAGTGCATGAAGCCCAAGGGACTTTGGTCCAGCATCTCAGCACTT 3099
QY 1918 CTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGGCCATCTCGCCAAACAAAGTGCCCTAC 1977
Db 3100 CTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGGCCATCTCGCCAAACAAAGTGCCCTAC 3159
QY 1978 TATATCAACCAGAGACTCAAACA 2001

Db 3160 TATATCAACCAGAGACTCAAACA 3183
RESULT 9
AAD37255
ID AAD37255 standard; DNA; 5149 BP.
XX
AC AAD37255;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-delta4173.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Unidentified.
XX
FN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 57-59; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence.
SQ Sequence 5149 BP; 1489 A; 1236 C; 1269 G; 1155 T; 0 other;
Query Match 83.1%; Score 1662.8; DB 24; Length 5149;
Best Local Similarity 88.3%; Pred.No. 0;
Matches 1929; Conservative 0; Mismatches 72; Indels 183; Gaps 4;
QY 1 GGCAGTTCAATGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 60
Db 1757 GGCAGTTCAATGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 1816
QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCAAGGAGAGATTTCTAAT 120
Db 1817 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCAAGGAGAGATTTCTAAT 1876
QY 121 GATGTGGAAGTGGTGAAGACCAAGTTTCATCTCATGAGGGGTACATGATGGATTGACA 180
Db 1877 GATGTGGAAGTGGTGAAGACCAAGTTTCATCTACTACTATGAGGGGTACATGATGGATTGACA 1936

QY 181 GCCATCAGGCGCGGTGGTAATATTCTACAAATTGGGAAGTAAGCTGATTGGAAACAGGA 240
Db 1937 GCCATCAGGCGCGGTGGTAATATTCTACAAATTGGGAAGTAAGCTGATTGGAAACAGGA 1996
QY 241 AAATTATCAGAAAGATGAAGAACTGAAGATCAAGAGCAGCATGAATCTCCTAAATCAAGA 300
Db 1997 AAATTATCAGAAAGATGAAGAACTGAAGATCAAGAGCAGCATGAATCTCCTAAATCAAGA 2056
QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATG 360
Db 2057 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATG 2116
QY 361 GATCTCCAGAATCAGAAACTGAAGAGTTGAATGACTGGCTFAACAAAAACAGAAAGA 420
Db 2117 GATCTCCAGAATCAGAAACTGAAGAGTTGAATGACTGGCTFAACAAAAACAGAAAGA 2176
QY 421 ACAAGGAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTA 480
Db 2177 ACAAGGAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTA 2236
QY 481 CAACAACATAAAGGTGCTTCAAGAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTC 540
Db 2237 CAACAACATAAAGGTGCTTCAAGAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTC 2296
QY 541 ACTCACATGTTGTTAGTTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAA 600
Db 2297 ACTCACATGTTGTTAGTTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAA 2356
QY 601 GAACAACITTAAGGTATTGGGAGATCGATGGCAACACATCTGTAGATGGACAGAACCCGC 660
Db 2357 GAACAACITTAAGGTATTGGGAGATCGATGGCAACACATCTGTAGATGGACAGAACCCGC 2416
QY 661 TGGGTTCTTTTACAAGACCAGCCTGAC-----CTAGCTCCTGGACTGACCACT 708
Db 2417 TGGGTTCTTTTACAAGACATCCTTCTCAATGGCAACGTCTTACTGAAGAACAGTGCCTT 2476
QY 709 ATTGGAGCCTCTCCTACTCAGACTGTTACTCTGGTGACACA-----ACCTGTGGTT 759
Db 2477 TTTAGTGCAATGCTTTTCAGAAAAAGAAAGATGCAGTGAACAAGATTACACAACACTGGCTTT 2536
QY 760 ACTAAGGAACTGCCATCTC-----779
Db 2537 AAAGATCAAAATCAAAATGTTATCAAGTCTTCAAAAACCTGGCCGTTTTAAAAGCGGATCTA 2596
QY 780 -----CAAAGTAAAGTCCATCTTCC 801
Db 2597 GAAAAGAAAAGCAATCCATGGGCNAAGTATTTCACTCAAAACAAGATCTTCTTTCAACA 2656
QY 802 TTGATGTTGGAG-----813
Db 2657 CTGAAGAATAAAGTCAGTCAGTGACCCAGAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGT 2716
QY 814 -----GTACCTACTCATAGATTACTGCAA 837
Db 2717 TGGGATAATTTAGTCCAAAAACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAA 2776
QY 838 CAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCCTGGCTTACAGAAGCTGAACAACACTGCC 897
Db 2777 CAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCCTGGCTTACAGAAGCTGAACAACACTGCC 2836
QY 898 AATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAGACTCCAAGGGAGTAAAA 957
Db 2837 AATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAGACTCCAAGGGAGTAAAA 2896
QY 958 GAGCTGATGAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTAT 1017
Db 2897 GAGCTGATGAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTAT 2956
QY 1018 CACAACCTGGATGAACACAGCCAAAAATCCTTGATCCCTGGAGGTTCCGATGATGCA 1077
Db 2957 CACAACCTGGATGAACACAGCCAAAAATCCTTGATCCCTGGAGGTTCCGATGATGCA 3016
QY 1078 GTCCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAG 1137

Db 3017 GTCCTGTTACAAAGACGTTTGGATAAACATGAACCTTCAAGTGGAGTGAACCTTCGGAATAAG 3076
QY 1138 TCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACAGTGAAGCGTCTGCACCTT 1197
Db 3077 TCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACAGTGAAGCGTCTGCACCTT 3136
QY 1198 TCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGAATTAAGCCGGCAGGCA 1257
Db 3137 TCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGAATTAAGCCGGCAGGCA 3196
QY 1258 CCTATTGGAGCGGACTTTCAGCAGTTTCAAGAGCAGAACGATGTACATAGGGCCTTCAAG 1317
Db 3197 CCTATTGGAGCGGACTTTCAGCAGTTTCAAGAGCAGAACGATGTACATAGGGCCTTCAAG 3256
QY 1318 AGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTT 1377
Db 3257 AGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTT 3316
QY 1378 CTGACAGAGCAGCCTTTTGGAAAGACTAGAGAAACTCTACAGGAGCCCGAGAGAGCTGCCT 1437
Db 3317 CTGACAGAGCAGCCTTTTGGAAAGACTAGAGAAACTCTACAGGAGCCCGAGAGAGCTGCCT 3376
QY 1438 CCTGAGGAGAGAGCCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAAT 1497
Db 3377 CCTGAGGAGAGAGCCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAAT 3436
QY 1498 ACTGAGTGGGAAAAAATTGAACCTTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACC 1557
Db 3437 ACTGAGTGGGAAAAAATTGAACCTTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACC 3496
QY 1558 CTTGAAAGACTCCAGGAACTTCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGCGCCAA 1617
Db 3497 CTTGAAAGACTCCAGGAACTTCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGCGCCAA 3556
QY 1618 GCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAA 1677
Db 3557 GCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAA 3616
QY 1678 GATCACCTCGAGAAAAAGTCAAGGCACTTCGAGGAGAAAAATTGCGCCTCTGAAAGAGAACGTG 1737
Db 3617 GATCACCTCGAGAAAAAGTCAAGGCACTTCGAGGAGAAAAATTGCGCCTCTGAAAGAGAACGTG 3676
QY 1738 AGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCCGTAT 1797
Db 3677 AGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCCGTAT 3736
QY 1798 AACCTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAG 1857
Db 3737 AACCTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAG 3796
QY 1858 GACCGAGTCAGGCAGCTGCATGAAGCCCAAGGACTTTGGTCCAGCATCTCAGCACTTT 1917
Db 3797 GACCGAGTCAGGCAGCTGCATGAAGCCCAAGGACTTTGGTCCAGCATCTCAGCACTTT 3856
QY 1918 CTTTCCACGTCGTCTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAAACAAAGTGCCCTAC 1977
Db 3857 CTTTCCACGTCGTCTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAAACAAAGTGCCCTAC 3916
QY 1978 TATATCAACCACGAGACTCAAAACA 2001
Db 3917 TATATCAACCACGAGACTCAAAACA 3940

RESULT 10

ABK81999

ID ABK81999 standard; DNA; 5462 BP.

XX

AC ABK81999;

XX

DT 13-AUG-2002 (first entry)

XX

DE DNA encoding mini-dystrophin protein deltaR2-R21+H3.

XX Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX OS Synthetic.

XX WO200229056-A2.
XX 11-APR-2002.
XX 04-OCT-2001; 2001WO-US31126.
XX 06-OCT-2000; 2000US-238848P.
XX (UNMI) UNIV MICHIGAN.
XX Chamberlain JS, Harper SQ;
XX WPI; 2002-435334/46.

XX A composition for preparing therapeutic drugs, has a mini-dystrophin
PT peptide comprising a specific number of spectrin-like repeat domains,
PT or a nucleic acid sequence encoding the mini-dystrophin peptide -
XX Disclosure; Fig 14; 145pp; English.

XX The invention describes a composition comprising a mini-dystrophin
CC peptide comprising a spectrin-like repeat domain, where the domain
CC comprises n spectrin-like repeats, and contains no more than n
CC spectrin-like repeats, where n is an even number between 4-24, or a
CC nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin
CC peptide or the polynucleotide encoding it is useful as a medicament,
CC for preparing a drug for therapeutic application and in the preparation
CC of a composition for treatment of muscle disease, e.g. Duchenne's
CC muscular dystrophy (DMD). This sequence represents a mini-dystrophin
CC sequence of the invention.

SQ Sequence 5462 BP; 1668 A; 1225 C; 1212 G; 1357 T; 0 other;
Query Match 64.1%; Score 1283; DB 24; Length 5462;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1302; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

QY 691 GCTCCTGGACTGACCACTATTGGAGCCTCTCCTACTCAGACTGTACTCTGGTGACACAA 750
DB 1547 GCTCCTGGACTGACCACTATTGGAGCCTCTCCTACTCAGACTGTACTCTGGTGACACAA 1606
QY 751 CCTGTGGTTACTAAGGAACTGCCATCTCCAACTAGAAATGCCATCTTCTTGATGTTG 810
DB 1607 CCTGTGGTTACTAAGGAACTGCCATCTCCAACTAGAAATGCCATCTTCTTGATGTTG 1666
QY 811 GAGGTACCTACTCATAGATTACTGCAACAGTTCCCTCGACCTGGAAGTTTCTTGCC 870
DB 1667 GAG-----CATAGATTACTGCAACAGTTCCCTCGACCTGGAAGTTTCTTGCC 1717
QY 871 TGGCTTACAGAACTGAAACAACTGCCATGTCTTACAGGATGCTACCCGTAAGGAAAGG 930
DB 1718 TGGCTTACAGAACTGAAACAACTGCCATGTCTTACAGGATGCTACCCGTAAGGAAAGG 1777
QY 931 CTCCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAAATGCAAGACCTCCAAAGT 990
DB 1778 CTCCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAAATGCAAGACCTCCAAAGT 1837
QY 991 GAAATTGAAGCTCACAGATGTTTATCACAACCTGGATGAAACAGCCAAATAATCCTG 1050
DB 1838 GAAATTGAAGCTCACAGATGTTTATCACAACCTGGATGAAACAGCCAAATAATCCTG 1897
QY 1051 AGATCCCTGGAAGTTCCGATGATGCACTCTGTTACAAAGACGTTTGGATAACATGAAC 1110
DB 1898 AGATCCCTGGAAGTTCCGATGATGCACTCTGTTACAAAGACGTTTGGATAACATGAAC 1957
QY 1111 TTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGT 1170

DB 1958 TTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGT 2017
QY 1171 TCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTG 1230
DB 2018 TCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTG 2077
QY 1231 AAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAAGAAG 1290
DB 2078 AAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAAGAAG 2137
QY 1291 CAGAACGATGTACATAGGGCCTTCAAGAGGGAATTAAGAACTTAAGAACTTGAATCATG 1350
DB 2138 CAGAACGATGTACATAGGGCCTTCAAGAGGGAATTAAGAACTTAAGAACTTGAATCATG 2197
QY 1351 AGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAA 1410
DB 2198 AGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAA 2257
QY 1411 CTCTACCAGGAGCCCGCAGAGAGCTGCCCTCCTGAGGAGAGAGCCAGAAATGTCACCTCGGCTT 1470
DB 2258 CTCTACCAGGAGCCCGCAGAGAGCTGCCCTCCTGAGGAGAGAGCCAGAAATGTCACCTCGGCTT 2317
QY 1471 CTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCT 1530
DB 2318 CTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCT 2377
QY 1531 GACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCCACG 1590
DB 2378 GACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCCACG 2437
QY 1591 GATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGSCAGCCCGTG 1650
DB 2438 GATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGSCAGCCCGTG 2497
QY 1651 GGCATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGA 1710
DB 2498 GGCATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGA 2557
QY 1711 GAAATTCGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACC 1770
DB 2558 GAAATTCGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACC 2617
QY 1771 ACTTTGGGCATTGAGCTCTCAGCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGA 1830
DB 2618 ACTTTGGGCATTGAGCTCTCAGCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGA 2677
QY 1831 TGGAAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAAGCAGCTGCATGAAGCCACAGG 1890
DB 2678 TGGAAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAAGCAGCTGCATGAAGCCACAGG 2737
QY 1891 GACTTTGTGTCAGCATCTCAGCACTTTCTTTCCAGTCTGTCCAGGGTCCCTGGGAGAGA 1950
DB 2738 GACTTTGTGTCAGCATCTCAGCACTTTCTTTCCAGTCTGTCCAGGGTCCCTGGGAGAGA 2797
QY 1951 GCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAACA 2001
DB 2798 GCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAACA 2848

RESULT 11
AAD06794
ID AAD06794 standard; DNA; 5952 BP.
XX
AC AAD06794;
XX
DT 06-AUG-2001 (first entry)
XX Human dystrophin gene (Becker form).
DE Human; dystrophin; extein; intein; trans-splicing; gene therapy;
KW Duchenne muscular dystrophy; Becker muscular dystrophy; DMD; BMD; ds.
XX

Db 4026 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGC 4085

Qy 1953 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 2001

Db 4086 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 4134

RESULT 12

ABK82000

ID ABK82000 standard; DNA; 8689 BP.

XX AC ABK82000;

XX DT 13-AUG-2002 (first entry)

XX DNA encoding mini-dystrophin protein deltaH2-R19.

DE Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;

XX KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.

XX OS Homo sapiens.

OS Synthetic.

XX WO200229056-A2.

PN 11-APR-2002.

XX 04-OCT-2001; 2001WO-US31126.

PF 06-OCT-2000; 2000US-238848P.

XX (UNMI) UNIV MICHIGAN.

PA Chamberlain JS, Harper SQ;

PI WPI; 2002-435334/46.

DR A composition for preparing therapeutic drugs, has a mini-dystrophin

XX PT peptide comprising a specific number of spectrin-like repeat domains,

PT or a nucleic acid sequence encoding the mini-dystrophin peptide -

XX PS Disclosure; Fig 15; 145pp; English.

XX The invention describes a composition comprising a mini-dystrophin

CC peptide comprising a spectrin-like repeat domain, where the domain

CC comprises n spectrin-like repeats, and contains no more than n

CC spectrin-like encoding a mini-dystrophin peptide. The mini-dystrophin

CC nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin

CC peptide or the polynucleotide encoding it is useful as a medicament,

CC for preparing a drug for therapeutic application and in the preparation

CC of a composition for treatment of muscle disease, e.g. Duchenne's

CC muscular dystrophy (DMD). This sequence represents a mini-dystrophin

CC sequence of the invention.

XX SQ Sequence 8689 BP; 2721 A; 1804 C; 1861 G; 2303 T; 0 other;

Query Match 59.1%; Score 1182.6; DB 24; Length 8689;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 813 GGTACCTACTCATAGATTACTGCAACAGTTCCCTCGGACCTGGAAAAGTTTCTTGCCCTG 872

Db 2992 GGAAGAAACTCATAGATTACTGCAACAGTTCCCTCGGACCTGGAAAAGTTTCTTGCCCTG 3051

Qy 873 GCTTACAGAAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGGCT 932

Db 3052 GCTTACAGAAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGGCT 3111

Qy 933 CCTAGAAGACTCCCAAGGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGA 992

Db 3112 CCTAGAAGACTCCCAAGGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGA 3171

Qy 993 AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAATCCTGAG 1052

Db 3172 AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAATCCTGAG 3231

Qy 1053 ATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTACAAAGACGTTTTGGATAACATGAACCTT 1112

Db 3232 ATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTACAAAGACGTTTTGGATAACATGAACCTT 3291

Qy 1113 CAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCCATTTTGAAGCCAGTTT 1172

Db 3292 CAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCCATTTTGAAGCCAGTTT 3351

Qy 1173 TGACCAGTGAAGCGTCTGCACCTTTTCTCTGCAGGAACCTTCTGGTGTGGTACAGCTGAA 1232

Db 3352 TGACCAGTGAAGCGTCTGCACCTTTTCTCTGCAGGAACCTTCTGGTGTGGTACAGCTGAA 3411

Qy 1233 AGATGATGAATTAAAGCCGCGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAAGCA 1292

Db 3412 AGATGATGAATTAAAGCCGCGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAAGCA 3471

Qy 1293 GAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGAG 1352

Db 3472 GAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGAG 3531

Qy 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTTAGAGAAACT 1412

Db 3532 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTTAGAGAAACT 3591

Qy 1413 CTACAGGAGCCCCAGAGAGCTGCTCCTCTGAGGAGAGAGCCCCAGAAATGTCACTCGGCTTCT 1472

Db 3592 CTACAGGAGCCCCAGAGAGCTGCTCCTCTGAGGAGAGAGCCCCAGAAATGTCACTCGGCTTCT 3651

Qy 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCACCTCCGCTGA 1532

Db 3652 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCACCTCCGCTGA 3711

Qy 1533 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGA 1592

Db 3712 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGA 3771

Qy 1593 TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCCGTGG 1652

Db 3772 TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCCGTGG 3831

Qy 1653 CGATCTCCTCATTTGACTCTCTCCAGATCACTCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 1712

Db 3832 CGATCTCCTCATTTGACTCTCTCCAGATCACTCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 3891

Qy 1713 AATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC 1772

Db 3892 AATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC 3951

Qy 1773 TTTGGGCATTTCAGCTCTCACCCGTATACCTCAGCAGCTCTGGAAGACCTGAACACCAGATG 1832

Db 3952 TTTGGGCATTTCAGCTCTCACCCGTATACCTCAGCAGCTCTGGAAGACCTGAACACCAGATG 4011

Qy 1833 GAAGCTTCTGCAAGTGGCCGTTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGA 1892

Db 4012 GAAGCTTCTGCAAGTGGCCGTTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGA 4071

Qy 1893 CTTTGGTCCAGCATCTCAGCAGCTTTCTTTCCACGTCTGTCCAGGTCCTTGGGAGAGAGC 1952

Db 4072 CTTTGGTCCAGCATCTCAGCAGCTTTCTTTCCACGTCTGTCCAGGTCCTTGGGAGAGAGC 4131

Qy 1953 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCAACCCAGAGACTCAAACA 2001

Db 4132 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCAACCCAGAGACTCAAACA 4180

RESULT 13

AAD37229

ID AAD37229 standard; DNA; 11058 BP.

XX

IC AAD37229;
CX 21-AUG-2002 (first entry)
XT Human dystrophin protein coding sequence.
CX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
CW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
CW Becker muscular dystrophy; ds.
CX Homo sapiens.
CX WO200183695-A2.
CX 08-NOV-2001.
CX 27-APR-2001; 2001WO-US13677.
CX 28-APR-2000; 2000US-200777P.
CX (XIAO/) XIAO X.
CX Xiao X;
CX WPI; 2002-049342/06.
CX New dystrophin minigene for treating Duchenne or Becker muscular
CX dystrophy comprises an N-terminal domain or modified N-terminal domain,
CX rod repeats, H1 and H4 domains and a cysteine rich domain of a
CX dystrophin gene -
CX Example 1; Page 40-43; 71pp; English.
CX The present invention relates to an isolated nucleotide sequence encoding
CX a dystrophin minigene. The minigene comprises N-terminal or modified
CX N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CX domains and cysteine-rich domains of dystrophin or utrophin genes. The
CX invention also relates to a recombinant adeno-associated virus (AAV)
CX comprising dystrophin minigene operably linked to an expression control
CX element. The dystrophin minigene in operable linkage with an expression
CX control element, in a recombinant adeno-associated virus or retrovirus is
CX useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CX dystrophy (BMD) in a mammalian subject. The present sequence is human
CX dystrophin protein coding sequence.
CX Sequence 11058 BP; 3686 A; 2292 C; 2621 G; 2459 T; 0 other;
CX
CX Query Match 59.1%; Score 1182.6; DB 24; Length 11058;
CX Best Local Similarity 99.7%; Pred. No. 0;
CX Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCCTGGACCTGGAAGTTTCTTGCTG 872
DB 8052 GGAAGAACTCATAGATTACTGCAACAGTTCCTGGACCTGGAAGTTTCTTGCTG 8111
QY 873 GCTTACAGAGCTGAAACAACTGCCAATGCTTACAGGATGCTACCCGTAAGGAAAGCT 932
DB 8112 GCTTACAGAGCTGAAACAACTGCCAATGCTTACAGGATGCTACCCGTAAGGAAAGCT 8171
QY 933 CCTAGAAGACTCCAGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 992
DB 8172 CCTAGAAGACTCCAGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 8231
QY 993 AATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAACAGCAAAAAATCCTGAG 1052
DB 8232 AATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAACAGCAAAAAATCCTGAG 8291
QY 1053 ATCCCTGGAAGGTTCCGATGTCAGTCCCTGTTACAAAGACGTTTGGATAACATGAACCT 1112
DB 8292 ATCCCTGGAAGGTTCCGATGTCAGTCCCTGTTACAAAGACGTTTGGATAACATGAACCT 8351
QY 1113 CAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAAGCCAGTTC 1172

DB 8352 CAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC 8411
QY 1173 TGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGTACAGCTGAA 1232
DB 8412 TGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGTACAGCTGAA 8471
QY 1233 AGATGATGAATTAAAGCCGCGCAGGACCTTATTTGGAGGCGGACCTTTCCAGCAGTTTCAAGCA 1292
DB 8472 AGATGATGAATTAAAGCCGCGCAGGACCTTATTTGGAGGCGGACCTTTCCAGCAGTTTCAAGCA 8531
QY 1293 GAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACTAAAGAACCTTGAATCATGAG 1352
DB 8532 GAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACTAAAGAACCTTGAATCATGAG 8591
QY 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGCAGCCCTTTGGAGGACTAGAGAAACT 1412
DB 8592 TACTCTTGAGACTGTACGAATATTTCTGACAGCAGCCCTTTGGAGGACTAGAGAAACT 8651
QY 1413 CTACCAGGAGCCCGCAGAGAGCTGCTCTGAGGAGAGAGCCCGCAGAAATGTCTCGGCTTCT 1472
DB 8652 CTACCAGGAGCCCGCAGAGAGCTGCTCTGAGGAGAGAGCCCGCAGAAATGTCTCGGCTTCT 8711
QY 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACTCTGACTCCGCTGA 1532
DB 8712 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACTCTGACTCCGCTGA 8771
QY 1533 CTGCGAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGA 1592
DB 8772 CTGCGAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGA 8831
QY 1593 TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCTTGGAGCCCGTGGG 1652
DB 8832 TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCTTGGAGCCCGTGGG 8891
QY 1653 CGATCTCTCATTTGACTCTCTCAAGATCACCTCGAGAAAGTCAAGGCATTTCCGAGGAGA 1712
DB 8892 CGATCTCTCATTTGACTCTCTCAAGATCACCTCGAGAAAGTCAAGGCATTTCCGAGGAGA 8951
QY 1713 AATTGCGCTCTGAAAGAGAGACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCAC 1772
DB 8952 AATTGCGCTCTGAAAGAGAGACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCAC 9011
QY 1773 TTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACAGATG 1832
DB 9012 TTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACAGATG 9071
QY 1833 GAAGCTTCTGAGGTGGCCGCTCGAGGACCGAGTCAGGACCTGTCATGAAGCCACAGGGA 1892
DB 9072 GAAGCTTCTGAGGTGGCCGCTCGAGGACCGAGTCAGGACCTGTCATGAAGCCACAGGGA 9131
QY 1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGCTCCCTGGGAGAGAGC 1952
DB 9132 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGCTCCCTGGGAGAGAGC 9191
QY 1953 CATCTCGCCAAACAAAGTGCCTACTATATCAACCACGAGACTCAAACA 2001
DB 9192 CATCTCGCCAAACAAAGTGCCTACTATATCAACCACGAGACTCAAACA 9240
RESULT 14
ABK82005
ID ABK82005 standard; DNA; 11241 BP.
XX
AC ABK82005;
XX
DT 13-AUG-2002 (first entry)
XX
DE cDNA encoding human dystrophin, full length HDMD.
XX
KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
XX Duchenne's muscular dystrophy; DMD; dystrophin; human; ds.
OS Homo sapiens.

XX
PN WO200229056-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US311126.
XX
PR 06-OCT-2000; 2000US-238848P.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Chamberlain JS, Harper SQ;
XX
DR WPI; 2002-435334/46.
XX
XX
PT A composition for preparing therapeutic drugs, has a mini-dystrophin
PT peptide comprising a specific number of spectrin-like repeat domains,
PT or a nucleic acid sequence encoding the mini-dystrophin peptide -
XX
PS Example 2; Fig 23; 145pp; English.
XX
CC The invention describes a composition comprising a mini-dystrophin
CC peptide comprising a spectrin-like repeat domain, where the domain
CC comprises n spectrin-like repeats, and contains no more than n
CC spectrin-like repeats, where n is an even number between 4-24, or a
CC nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin
CC peptide or the polynucleotide encoding it is useful as a medicament,
CC for preparing a drug for therapeutic application and in the preparation
CC of a composition for treatment of muscle disease, e.g. Duchenne's
CC muscular dystrophy (DMD). This sequence represents a human dystrophin
CC polynucleotide sequence used in the creation of the mini-dystrophin
CC peptides of the invention.
XX
SQ Sequence 11241 BP; 3738 A; 2325 C; 2656 G; 2522 T; 0 other;
Query Match 59.1%; Score 1182.6; DB 24; Length 11241;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 813 GGTACTACTCATAGATTACTGCAACAGTTCCCTGGACCTGGAAAGTTTCTTGCCTG 872
Db 8260 GGAAGAACTCATAGATTACTGCAACAGTTCCCTGGACCTGGAAAGTTTCTTGCCTG 8319
QY 873 GCTTACAGAGCTGAAACAACTGCCAATGTCTCTACAGATGCTACCCGTAGGAAAGGCT 932
Db 8320 GCTTACAGAGCTGAAACAACTGCCAATGTCTCTACAGATGCTACCCGTAGGAAAGGCT 8379
QY 933 CCTAGAGACTCCAAGGGAGTAAAGAGCTGTATGAAACAAATGGCAAGACCTCCAAGGTGA 992
Db 8380 CCTAGAGACTCCAAGGGAGTAAAGAGCTGTATGAAACAAATGGCAAGACCTCCAAGGTGA 8439
QY 993 AATTGAAGCTCACACAGATGTTTATCAAACTGGATGAAACAGCCCAAAAATCTCTGAG 1052
Db 8440 AATTGAAGCTCACACAGATGTTTATCAAACTGGATGAAACAGCCCAAAAATCTCTGAG 8499
QY 1053 ATCCCTGGAAGTTCGGATGATGCACTCTCTTACAAAGACGTTTGGATAACATGAACCTT 1112
Db 8500 ATCCCTGGAAGTTCGGATGATGCACTCTCTTACAAAGACGTTTGGATAACATGAACCTT 8559
QY 1113 CAAGTGGAGTGAACCTCGGAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTC 1172
Db 8560 CAAGTGGAGTGAACCTCGGAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTC 8619
QY 1173 TGACCAGTGAAGCGTCTGCACCTTTCTCTGAGGAACCTTCTGTGTGGCTACAGCTGAA 1232
Db 8620 TGACCAGTGAAGCGTCTGCACCTTTCTCTGAGGAACCTTCTGTGTGGCTACAGCTGAA 8679
QY 1233 AGATGATGAATTAAGCCGGCAGGCACCTTATTGGAGGCGACTTTCAGAGCAGTTTCAGAAACA 1292
Db 8680 AGATGATGAATTAAGCCGGCAGGCACCTTATTGGAGGCGACTTTCAGAGCAGTTTCAGAAACA 8739
QY 1293 GAACGATGTACATAGGCGCTTCAAGAGGGGAATTGAAACTAAAGAACCTGTATCATGAG 1352
|||

Db 8740 GAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACCTAAGAACCTGTATCATGAG 8799
QY 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT 1412
Db 8800 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT 8859
QY 1413 CTACCAGGAGCCCAAGAGAGTGCCTCTCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCT 1472
Db 8860 CTACCAGGAGCCCAAGAGAGTGCCTCTCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCT 8919
QY 1473 ACGAAAGCAGGCTGAGGAGGTCAATCTGAGTGGGAAAAAATTGAACCTGCACTCCGCTGA 1532
Db 8920 ACGAAAGCAGGCTGAGGAGGTCAATCTGAGTGGGAAAAAATTGAACCTGCACTCCGCTGA 8979
QY 1533 CTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGA 1592
Db 8980 CTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGA 9039
QY 1593 TGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 1652
Db 9040 TGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 9099
QY 1653 CGATCTCTCATTTGACTCTCTCCAAGATCACTCGAGAAAGTCAAGGCACTTCGAGGAGA 1712
Db 9100 CGATCTCTCATTTGACTCTCTCCAAGATCACTCGAGAAAGTCAAGGCACTTCGAGGAGA 9159
QY 1713 AATTGCGCTCTGAAAGAGAAACGTGAGCCACGTCATGACCTTGTCTGCGCAGCTTACCAC 1772
Db 9160 AATTGCGCTCTGAAAGAGAAACGTGAGCCACGTCATGACCTTGTCTGCGCAGCTTACCAC 9219
QY 1773 TTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACAGATG 1832
Db 9220 TTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACAGATG 9279
QY 1833 GAAGTTCTGAGGTGGCGCTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGA 1892
Db 9280 GAAGTTCTGAGGTGGCGCTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGA 9339
QY 1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCGTCCAGGGTCCCTGGGAGAGAGC 1952
Db 9340 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCGTCCAGGGTCCCTGGGAGAGAGC 9399
QY 1953 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 2001
Db 9400 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 9448
RESULT 15
ABK82002
ID ABK82002 standard; DNA; 11443 BP.
XX
AC ABK82002;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding mini-dystrophin protein deltaR9-R16.
XX
KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200229056-A2.
XX
XX 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US311126.
XX
PR 06-OCT-2000; 2000US-238848P.
XX
PA (UNMI) UNIV MICHIGAN.
XX

PI Chamberlain JS, Harper SQ;

XX WPI; 2002-435334/46.

DR A composition for preparing therapeutic drugs, has a mini-dystrophin
XX peptide comprising a specific number of spectrin-like repeat domains,
PT or a nucleic acid sequence encoding the mini-dystrophin peptide -

PT Disclosure; Fig 17; 145pp; English.

XX The invention describes a composition comprising a mini-dystrophin
XX peptide comprising a spectrin-like repeat domain, where the domain
XX comprises n spectrin-like repeats, and contains no more than n
XX spectrin-like repeats, where n is an even number between 4-24, or a
XX nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin
XX peptide or the polynucleotide encoding it is useful as a medicament,
XX for preparing a drug for therapeutic application and in the preparation
XX of a composition for treatment of muscle disease, e.g. Duchenne's
XX muscular dystrophy (DMD). This sequence represents a mini-dystrophin
XX sequence of the invention.

SQ Sequence 11443 BP; 3707 A; 2339 C; 2502 G; 2895 T; 0 other;

Query Match 59.1%; Score 1182.6; DB 24; Length 11443;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCCCTCGGACCTGGAAAAGTTTCTTGCCCTG 872

Db 5746 GGAAGAACTCATAGATTACTGCAACAGTTCCCTCGGACCTGGAAAAGTTTCTTGCCCTG 5805

QY 873 GCTTACAGAACTGAAACAACTGCCAATGTCTACAGGATGCTACCCGTAAGGAAAGGCT 932

Db 5806 GCTTACAGAACTGAAACAACTGCCAATGTCTACAGGATGCTACCCGTAAGGAAAGGCT 5865

QY 933 CCTAGAAAGACTCCAAGGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAGGTGA 992

Db 5866 CCTAGAAAGACTCCAAGGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAGGTGA 5925

QY 993 AATTGAAGCTCACACAGATGTTTATCAACACCTGGATGAAAACAGCCAAAATCCTGAG 1052

Db 5926 AATTGAAGCTCACACAGATGTTTATCAACACCTGGATGAAAACAGCCAAAATCCTGAG 5985

QY 1053 ATCCCTGGAAGTCCGATGATGCGAGTCCTGTTACAAAGACGTTTGGATAACATGAACCTT 1112

Db 5986 ATCCCTGGAAGTCCGATGATGCGAGTCCTGTTACAAAGACGTTTGGATAACATGAACCTT 6045

QY 1113 CAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAAGCCAGTTC 1172

Db 6046 CAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAAGCCAGTTC 6105

QY 1173 TGACCAGTGGAGCGTCTGCACCTTTCTGACAGGAACTTCTGGTGTGGCTACAGCTGAA 1232

Db 6106 TGACCAGTGGAGCGTCTGCACCTTTCTGACAGGAACTTCTGGTGTGGCTACAGCTGAA 6165

QY 1233 AGATGATGAATTAAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAAACA 1292

Db 6166 AGATGATGAATTAAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAAACA 6225

QY 1293 GAACGATGTACATAGGGCCCTTCAAGAGGGAATGAAAACCTAAGAACCTGTAAATCATGAG 1352

Db 6226 GAACGATGTACATAGGGCCCTTCAAGAGGGAATGAAAACCTAAGAACCTGTAAATCATGAG 6285

QY 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACT 1412

Db 6286 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACT 6345

QY 1413 CTACCAGGAGCCAGAGAGCTGCCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT 1472

Db 6346 CTACCAGGAGCCAGAGAGCTGCCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT 6405

QY 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGAACCTGCACCTCCGCTGA 1532

Db 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGAACCTGCACCTCCGCTGA 1532

Db 6406 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGAACCTGCACCTCCGCTGA 6465

QY 1533 CTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGA 1592

Db 6466 CTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGA 6525

QY 1593 TGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 1652

Db 6526 TGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 6585

QY 1653 CGATCTCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCATTTTCGAGGAGA 1712

Db 6586 CGATCTCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCATTTTCGAGGAGA 6645

QY 1713 AATTGGCCCTCTGAAAAGAGAACTGAGCCACGTCATCAATGACCTTGTCTGCCAGCTTACCAC 1772

Db 6646 AATTGGCCCTCTGAAAAGAGAACTGAGCCACGTCATCAATGACCTTGTCTGCCAGCTTACCAC 6705

QY 1773 TTTGGSCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATG 1832

Db 6706 TTTGGSCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATG 6765

QY 1833 GAAGCTTCTGCAAGTGGCCGTCGAGGACCGAGTCAGGCAAGTGCATGAAGCCACAGGGA 1892

Db 6766 GAAGCTTCTGCAAGTGGCCGTCGAGGACCGAGTCAGGCAAGTGCATGAAGCCACAGGGA 6825

QY 1893 CTTTGGTCCAGCATCTCAGCACCTTTCTTTCCACGTCCTGCCAGGGTCCCTGGGAGAGAGC 1952

Db 6826 CTTTGGTCCAGCATCTCAGCACCTTTCTTTCCACGTCCTGCCAGGGTCCCTGGGAGAGAGC 6885

QY 1953 CATCTCGCCAAACAAAGTGCCCTACTATATATCAACCACGAGACTCAAACA 2001

Db 6886 CATCTCGCCAAACAAAGTGCCCTACTATATATCAACCACGAGACTCAAACA 6934

Search completed: February 1, 2004, 12:04:09

Job time : 487.292 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

AM nucleic - nucleic search, using sw model
run on: February 1, 2004, 10:14:55 ; Search time 7128.21 Seconds
(without alignments)
11483.979 Million cell updates/sec
Title: US-09-845-416-6_COPY_1000_3000
Perfect score: 2001
Sequence: 1 ggcagttcattgatggagag.....tcaaccacgagactcaaac 2001
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|--------|---------------|--------|-------|---------------------|
| 1 | 1283 | 64.1 | 5462 | 6 | AX538621 Sequence |
| 2 | 1182.6 | 59.1 | 5952 | 6 | AR304538 Sequence |
| 3 | 1182.6 | 59.1 | 5952 | 6 | AX114289 Sequence |
| 4 | 1182.6 | 59.1 | 8689 | 6 | AX538622 Sequence |
| 5 | 1182.6 | 59.1 | 11443 | 6 | AX538624 Sequence |
| 6 | 1182.6 | 59.1 | 12057 | 6 | AX538627 Sequence |
| 7 | 1182.6 | 59.1 | 12446 | 9 | X14298 Human mRNA |
| 8 | 1182.6 | 59.1 | 13957 | 6 | AX409637 Sequence |
| 9 | 1182.6 | 59.1 | 13957 | 6 | AX538581 Sequence |
| 10 | 1182.6 | 59.1 | 13957 | 9 | M18533 Homo sapien |
| 11 | 1182.6 | 59.1 | 13977 | 6 | AR220819 Sequence |
| 12 | 1180 | 59.0 | 5339 | 6 | AX538620 Sequence |
| 13 | 1030.6 | 51.5 | 13887 | 4 | AF070485 Canis fam |
| 14 | 1004 | 50.2 | 13815 | 6 | AX306153 Sequence |
| 15 | 1004 | 50.2 | 13815 | 6 | AX538582 Sequence |
| 16 | 1004 | 50.2 | 13815 | 10 | M68859 Mouse dyster |
| 17 | 1004 | 50.2 | 19307 | 6 | AR093392 Sequence |
| 18 | 1004 | 50.2 | 19307 | 6 | AR142592 Sequence |
| 19 | 765 | 38.2 | 13575 | 5 | GGDYS |
| 20 | 727 | 36.3 | 5417 | 6 | AX538619 |
| 21 | 678.4 | 33.9 | 4402 | 6 | E30220 |
| 22 | 618.8 | 30.9 | 4075 | 6 | E30221 |
| 23 | 614.2 | 30.7 | 4402 | 6 | E30219 |
| 24 | 547.4 | 27.4 | 3275 | 10 | MUSDYS |
| 25 | 459.4 | 23.0 | 630 | 9 | HSDMDF1 |
| 26 | 422.6 | 21.1 | 3521 | 5 | AF339031 |
| 27 | 408.2 | 20.4 | 11096 | 6 | AX538584 |
| 28 | 408.2 | 20.4 | 11096 | 10 | MMY12229 |
| 29 | 408 | 20.4 | 10705 | 10 | RNAJ2967 |
| 30 | 397 | 19.8 | 6045 | 6 | A63605 |
| 31 | 397 | 19.8 | 6045 | 6 | AR281528 |
| 32 | 397 | 19.8 | 6059 | 6 | AX107972 |
| 33 | 397 | 19.8 | 10302 | 6 | AX538583 |
| 34 | 397 | 19.8 | 10302 | 9 | HSMUPS |
| 35 | 397 | 19.8 | 10320 | 6 | A63607 |
| 36 | 397 | 19.8 | 10320 | 6 | AR281529 |
| 37 | 391.8 | 19.6 | 2654 | 5 | FSCDYSTRO |
| 38 | 387 | 19.3 | 387 | 6 | AX538612 |
| 39 | 378.6 | 18.9 | 3161 | 10 | MMGUTRPH |
| 40 | 348 | 17.4 | 348 | 6 | AX538611 |
| 41 | 330 | 16.5 | 333 | 6 | AX538589 |
| 42 | 327 | 16.3 | 327 | 6 | AX538588 |
| 43 | 324 | 16.2 | 324 | 6 | AX538613 |
| 44 | 322.6 | 16.1 | 3499 | 9 | HSU43519 |
| 45 | 322.6 | 16.1 | 5106 | 6 | AX552248 Sequence |

ALIGNMENTS

| | | | | | |
|------------|--|-------------|-----|--------|-----------------|
| RESULT 1 | | | | | |
| AX538621 | | | | | |
| LOCUS | AX538621 | 5462 bp | DNA | linear | PAT 23-NOV-2002 |
| DEFINITION | Sequence 41 from Patent WO0229056. | | | | |
| ACCESSION | AX538621 | | | | |
| VERSION | AX538621.1 | GI:25271168 | | | |
| KEYWORDS | | | | | |
| SOURCE | synthetic construct | | | | |
| ORGANISM | synthetic construct | | | | |
| REFERENCE | artificial sequences. | | | | |
| AUTHORS | 1 | | | | |
| TITLE | Chamberlain, J.S. and Harper, S.Q. | | | | |
| JOURNAL | Mini-dystrophin nucleic acid and peptide sequences | | | | |
| | Patent: WO 0229056-A 41 11-APR-2002; | | | | |
| | THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US) | | | | |

| | | | |
|-----------------------|-----------------|--|--------------------|
| FEATURES | | Location/Qualifiers | |
| source | 1..5462 | /organism="synthetic construct" | |
| | | /mol_type="genomic DNA" | |
| | | /db_xref="taxon:32630" | |
| | | /note="Synthetic" | |
| BASE COUNT | 1668 a | 1225 c | 1212 g 1357 t |
| ORIGIN | | | |
| Query Match | 64.1%; | Score 1283; | DB 6; Length 5462; |
| Best Local Similarity | 99.3%; | Pred. No. 4.5e-311; | |
| Matches 1302; | Conservative 0; | Mismatches 0; | Indels 9; Gaps 1; |
| QY | 691 | GCTCCTGGACTGACCACCTATTGGAGCCTCTCCTACTCAGACTGTTACTCTGGTGACACAA | 750 |
| Db | 1547 | GCTCCTGGACTGACCACCTATTGGAGCCTCTCCTACTCAGACTGTTACTCTGGTGACACAA | 1606 |
| QY | 751 | CCTGTGGTTACTAAGGAAACTGCCATCTCCAACTAGAAATGCCATCTTCCTTGATGTTG | 810 |
| Db | 1607 | CCTGTGGTTACTAAGGAAACTGCCATCTCCAACTAGAAATGCCATCTTCCTTGATGTTG | 1666 |
| QY | 811 | GAGTACCTACTCATAGATTACTGCAACAGTTTCCCTCGGACCTGGAAAAGTTTCTTGCC | 870 |
| Db | 1667 | GAG-----CATAGATTACTGCAACAGTTTCCCTCGGACCTGGAAAAGTTTCTTGCC | 1717 |
| QY | 871 | TGGCTTACAGAAGCTGAAACRACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGG | 930 |
| Db | 1718 | TGGCTTACAGAAGCTGAAACRACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGG | 1777 |
| QY | 931 | CTCCTAGAAGACTCCAAGGGAGTAAAGAGCTGTATGAACAATGGCAAGACCTTCCAAGGT | 990 |
| Db | 1778 | CTCCTAGAAGACTCCAAGGGAGTAAAGAGCTGTATGAACAATGGCAAGACCTTCCAAGGT | 1837 |
| QY | 991 | GAAATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAACAGCCAAAATAATCCTG | 1050 |
| Db | 1838 | GAAATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAACAGCCAAAATAATCCTG | 1897 |
| QY | 1051 | AGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTACAAAGACGTTTGGATAACATGAAC | 1110 |
| Db | 1898 | AGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTACAAAGACGTTTGGATAACATGAAC | 1957 |
| QY | 1111 | TTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGTCCCATTTTGGAGCCAGT | 1170 |
| Db | 1958 | TTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGTCCCATTTTGGAGCCAGT | 2017 |
| QY | 1171 | TCTGACCAGTGAAGCGTCTGCACCTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTG | 1230 |
| Db | 2018 | TCTGACCAGTGAAGCGTCTGCACCTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTG | 2077 |
| QY | 1231 | AAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGGGAGCTTTCAGCAGGTTTCAGAAG | 1290 |
| Db | 2078 | AAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGGGAGCTTTCAGCAGGTTTCAGAAG | 2137 |
| QY | 1291 | CAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTATCATG | 1350 |
| Db | 2138 | CAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTATCATG | 2197 |
| QY | 1351 | AGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAA | 1410 |
| Db | 2198 | AGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAA | 2257 |
| QY | 1411 | CTCTACCAGGAGCCACAGAGAGCTGCCTCTGAGGAGAGAGCCCAAGAATGTCACTCGGCTT | 1470 |
| Db | 2258 | CTCTACCAGGAGCCACAGAGAGCTGCCTCTGAGGAGAGAGCCCAAGAATGTCACTCGGCTT | 2317 |
| QY | 1471 | CTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACCTCCGCT | 1530 |
| Db | 2318 | CTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACCTCCGCT | 2377 |
| QY | 1531 | GACTGGCAGAGAAAAATAGATGAGACCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACG | 1590 |
| Db | 2378 | GACTGGCAGAGAAAAATAGATGAGACCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACG | 2437 |

| | | | |
|-----------------------|---|---|----------------------------|
| QY | 1591 | GATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTG | 1650 |
| Db | 2438 | GATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTG | 2497 |
| QY | 1651 | GGCGATCTCCTCATGACTCTCTCTCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGA | 1710 |
| Db | 2498 | GGCGATCTCCTCATGACTCTCTCTCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGA | 2557 |
| QY | 1711 | GAAATTGCGCCTCTGAAAGAGAACTGAGGACCGTCAATGACCTTGCTCGCCAGCTTACC | 1770 |
| Db | 2558 | GAAATTGCGCCTCTGAAAGAGAACTGAGGACCGTCAATGACCTTGCTCGCCAGCTTACC | 2617 |
| QY | 1771 | ACTTTGGGCATTACGCTCTCACCGTATAAACCCTAGCACTCTGGAAGACCTGAACACCAGA | 1830 |
| Db | 2618 | ACTTTGGGCATTACGCTCTCACCGTATAAACCCTAGCACTCTGGAAGACCTGAACACCAGA | 2677 |
| QY | 1831 | TGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGG | 1890 |
| Db | 2678 | TGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGG | 2737 |
| QY | 1891 | GACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCGTCCAGGGTCCCTGGGAGAGA | 1950 |
| Db | 2738 | GACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCGTCCAGGGTCCCTGGGAGAGA | 2797 |
| QY | 1951 | GCCATCTCGCCAAACAAAGTGCCTACTATATCAACACGAGACTCAAAACA | 2001 |
| Db | 2798 | GCCATCTCGCCAAACAAAGTGCCTACTATATCAACACGAGACTCAAAACA | 2848 |
| RESULT 2 | | | |
| AR304538 | AR304538 | 5952 bp | DNA linear PAT 12-JUN-2003 |
| LOCUS | Sequence 1 from patent US 6544786. | | |
| DEFINITION | AR304538 | | |
| ACCESSION | AR304538.1 | GI:31693691 | |
| VERSION | | | |
| KEYWORDS | Unknown. | | |
| SOURCE | Unknown. | | |
| ORGANISM | Unclassified. | | |
| REFERENCE | 1 (bases 1 to 5952) | | |
| AUTHORS | Xiao,X. and Liu,P.X. | | |
| TITLE | Method and vector for producing and transferring trans-spliced peptides | | |
| JOURNAL | Patent: US 6544786-A 1 08-APR-2003; | | |
| FEATURES | Location/Qualifiers | | |
| source | 1..5952 | | |
| BASE COUNT | 1860 a | 1344 c | 1410 g 1338 t |
| ORIGIN | | | |
| Query Match | 59.1%; | Score 1182.6; | DB 6; Length 5952; |
| Best Local Similarity | 99.7%; | Pred. No. 7.6e-286; | |
| Matches 1185; | Conservative 0; | Mismatches 4; | Indels 0; Gaps 0; |
| QY | 813 | GGTACCTACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGAAAAGTTTCTTGCCTG | 872 |
| Db | 2946 | GGAAGAACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGAAAAGTTTCTTGCCTG | 3005 |
| QY | 873 | GCTTACAGAACTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGGCT | 932 |
| Db | 3006 | GCTTACAGAACTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGGCT | 3065 |
| QY | 933 | CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA | 992 |
| Db | 3066 | CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA | 3125 |
| QY | 993 | AATTGAAGCTCACACAGATGTTTATCAACACCTGGATGAAACAGCCAAAATAATCCTGAG | 1052 |
| Db | 3126 | AATTGAAGCTCACACAGATGTTTATCAACACCTGGATGAAACAGCCAAAATAATCCTGAG | 3185 |
| QY | 1053 | ATCCCTGGAAGTTCCGATGATGCAGTCTGTACAAAGACGTTTGGATAACATGAACCTT | 1112 |
| Db | 3186 | ATCCCTGGAAGTTCCGATGATGCAGTCTGTACAAAGACGTTTGGATAACATGAACCTT | 3245 |

QY 1113 CAAGTGGAGTGAACCTTCGGAAAAGTCTCTCAACATTAGGTCCCATTTGGAGCCAGTTC 1172
Db |||||
3246 CAAGTGGAGTGAACCTTCGGAAAAGTCTCTCAACATTAGGTCCCATTTGGAGCCAGTTC 3305
QY 1173 TGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 1232
Db |||||
3306 TGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 3365
QY 1233 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTTCAGAAGCA 1292
Db |||||
3366 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTTCAGAAGCA 3425
QY 1293 GAACGATGTACATAGGGCCTTCAAGAGGGAATTAAGAACTTAAAGAACTTGAATCATGAG 1352
Db |||||
3426 GAACGATGTACATAGGGCCTTCAAGAGGGAATTAAGAACTTAAAGAACTTGAATCATGAG 3485
QY 1353 TACTCTTGAGACTGTACGAATATTCTTGACAGAGCAGCCCTTTGGAAAGGACTAGAGAACT 1412
Db |||||
3486 TACTCTTGAGACTGTACGAATATTCTTGACAGAGCAGCCCTTTGGAAAGGACTAGAGAACT 3545
QY 1413 CTACCAGGAGCCAGAGAGCTGCCTCTTGAGGAGAGAGCCAGAACTGTCACTCGGCTTCT 1472
Db |||||
3546 CTACCAGGAGCCAGAGAGCTGCCTCTTGAGGAGAGAGCCAGAACTGTCACTCGGCTTCT 3605
QY 1473 ACGAAAGCAGCTGAGGAGGTCAATACTAGTGGGAAAATTAAGAACTTGAATCATGAG 1532
Db |||||
3606 ACGAAAGCAGCTGAGGAGGTCAATACTAGTGGGAAAATTAAGAACTTGAATCATGAG 3665
QY 1533 CTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGA 1592
Db |||||
3666 CTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGA 3725
QY 1593 TGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCTGGCAGCCCGTGGG 1652
Db |||||
3726 TGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCTGGCAGCCCGTGGG 3785
QY 1653 CGATCTCCTCATTTGACTCTCTCCAAGATCACTCGAGAAAGTCAAGGCACTTCGAGGAGA 1712
Db |||||
3786 CGATCTCCTCATTTGACTCTCTCCAAGATCACTCGAGAAAGTCAAGGCACTTCGAGGAGA 3845
QY 1713 AATTGGCCCTCTGAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC 1772
Db |||||
3846 AATTGGCCCTCTGAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC 3905
QY 1773 TTTGGGCATTGAGCTCTACCGGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATG 1832
Db |||||
3906 TTTGGGCATTGAGCTCTACCGGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATG 3965
QY 1833 GAAGCTTCTGAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCACACAGGGA 1892
Db |||||
3966 GAAGCTTCTGAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCACACAGGGA 4025
QY 1893 CTTTGGTCCAGCATCTCAGCACTTTCTTCCAGTCTGTCCAGGGTCCCTGGGAGAGAGC 1952
Db |||||
4026 CTTTGGTCCAGCATCTCAGCACTTTCTTCCAGTCTGTCCAGGGTCCCTGGGAGAGAGC 4085
QY 1953 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAACA 2001
Db |||||
4086 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAACA 4134

RESULT 3
AX114289
LOCUS AX114289 5952 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 1 from Patent WO0129243.
ACCESSION AX114289
VERSION AX114289.1 GI:14031259
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Paul,X.L. and Xiao,X.
TITLE Method and vector for producing and transferring trans-spliced peptides
JOURNAL Patent: WO 0129243-A 1 26-APR-2001;
DALHOUSIE UNIVERSITY (CA) ; UNIV. OF PITTSBURGH OF THE COMMONWEALTH
SYSTEM OF HIGHER EDUCATION (US)
FEATURES Location/Qualifiers
source 1..5952
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
misc_feature 2897..2898
/note="S4 junction site"
misc_feature 3198..3199
/note="S2 junction site"
BASE COUNT 1860 a 1344 c 1410 g 1338 t
ORIGIN
Query Match 59.1%; Score 1182.6; DB 6; Length 5952;
Best Local Similarity 99.7%; Pred. No. 7.6e-286;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAGTTTCTTGCCTG 872
Db |||||
2946 GGAAGAAACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAGTTTCTTGCCTG 3005
QY 873 GCTTACAGAGCTGAAACAACTGCCAATGTCTACAGGATGCTACCCGTAAGGAAAGGCT 932
Db |||||
3006 GCTTACAGAGCTGAAACAACTGCCAATGTCTACAGGATGCTACCCGTAAGGAAAGGCT 3065
QY 933 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 992
Db |||||
3066 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 3125
QY 993 AATTGAAGCTCACACAGATGTTTATCAACCTTGGATGAAACAGCCCAAAAATCCTGAG 1052
Db |||||
3126 AATTGAAGCTCACACAGATGTTTATCAACCTTGGATGAAACAGCCCAAAAATCCTGAG 3185
QY 1053 ATCCCTGGAAGTTCGGATGATGCAGTCTGTTACAAAGACGTTTGGATAACATGAACCTT 1112
Db |||||
3186 ATCCCTGGAAGTTCGGATGATGCAGTCTGTTACAAAGACGTTTGGATAACATGAACCTT 3245
QY 1113 CAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC 1172
Db |||||
3246 CAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC 3305
QY 1173 TGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAA 1232
Db |||||
3306 TGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAA 3365
QY 1233 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTTCAGAAGCA 1292
Db |||||
3366 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTTCAGAAGCA 3425
QY 1293 GAACGATGTACATAGGGCCTTCAAGAGGGAATTAAGAACTTAAAGAACTTGAATCATGAG 1352
Db |||||
3426 GAACGATGTACATAGGGCCTTCAAGAGGGAATTAAGAACTTGAATCATGAG 3485
QY 1353 TACTCTTGAGACTGTACGAATATTCTTGACAGAGCAGCCCTTTGGAAAGGACTAGAGAACT 1412
Db |||||
3486 TACTCTTGAGACTGTACGAATATTCTTGACAGAGCAGCCCTTTGGAAAGGACTAGAGAACT 3545
QY 1413 CTACCAGGAGCCAGAGAGCTGCCTCTTGAGGAGAGAGCCAGAACTGTCACTCGGCTTCT 1472
Db |||||
3546 CTACCAGGAGCCAGAGAGCTGCCTCTTGAGGAGAGAGCCAGAACTGTCACTCGGCTTCT 3605
QY 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTAAGAACTTGAACCTGCACTCCGCTGA 1532
Db |||||
3606 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTAAGAACTTGAACCTGCACTCCGCTGA 3665
QY 1533 CTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGA 1592
Db |||||

Db 3666 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGA 3725

QY 1593 TGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGGAGCCCGTGGG 1652

Db 3726 TGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGGAGCCCGTGGG 3785

QY 1653 CGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 1712

Db 3786 CGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 3845

QY 1713 AATTGGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC 1772

Db 3846 AATTGGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC 3905

QY 1773 TTTGGGCAATTCAGCTCTCAACCGTATAACCTCAGCACTCTGGAAGACCTGGAACACCATG 1832

Db 3906 TTTGGGCAATTCAGCTCTCAACCGTATAACCTCAGCACTCTGGAAGACCTGGAACACCATG 3965

QY 1833 GAAGCTTCTGCAGGTGGCCGTCGAGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGA 1892

Db 3966 GAAGCTTCTGCAGGTGGCCGTCGAGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGA 4025

QY 1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTGTCTCAGGGTCCCTGGGAGAGAGC 1952

Db 4026 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTGTCTCAGGGTCCCTGGGAGAGAGC 4085

QY 1953 CATCTCGCCAAACAAAGTGCCCTACTATATCAACACAGAGACTCAAACA 2001

Db 4086 CATCTCGCCAAACAAAGTGCCCTACTATATCAACACAGAGACTCAAACA 4134

RESULT 4

AX538622

LOCUS

DEFINITION

AX538622

ACCESSION

VERSION

AX538622.1 GI:25271171

KEYWORDS

SOURCE

ORGANISM

artificial sequences.

REFERENCE

1

AUTHORS

Chamberlain, J.S. and Harper, S.Q.

TITLE

Mini-dystrophin nucleic acid and peptide sequences

JOURNAL

Patent: WO 0229056-A 42 11-APR-2002;

THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

FEATURES

Location/Qualifiers

1..8689

/organism="synthetic construct"

/mol_type="genomic DNA"

/db_xref="taxon:32630"

/note="Synthetic"

BASE COUNT 2721 a 1804 c 1861 g 2303 t

ORIGIN

Query Match 59.1%; Score 1182.6; DB 6; Length 8689;

Best Local Similarity 99.7%; Pred. No. 7.7e-286;

Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTG 872

Db 2992 GGAAGAACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTG 3051

QY 873 GCTTACAGAAGCTGAAACAACTGCAATGTCTTACAGGATGCTACCCGTAGGAAAGGCT 932

Db 3052 GCTTACAGAAGCTGAAACAACTGCAATGTCTTACAGGATGCTACCCGTAGGAAAGGCT 3111

QY 933 CCTAGAAGACTCCAAGCGAGTAAAGAGCTGATGAACAAATGGCAAGACCTCCAAGGTGA 992

Db 3112 CCTAGAAGACTCCAAGCGGAGTAAAGAGCTGATGAACAAATGGCAAGACCTCCAAGGTGA 3171

QY 993 AATTGAAGCTCACACAGATGTTTATCACAACTTGTGATGAACAGCCAAATAATCCTTGA 1052

Db 3172 AATTGAAGCTCACACAGATGTTTATCAACACCTGGATGAAACACGCCAAATAATCCTGAG 3231

QY 1053 ATCCCTGGAAGGTTCCGATGATGCAGTCTGTTCACAAAGACGTTTGGATTAACATGAACCTT 1112

Db 3232 ATCCCTGGAAGGTTCCGATGATGCAGTCTGTTCACAAAGACGTTTGGATTAACATGAACCTT 3291

QY 1113 CAAGTGGAGTGAACCTTCGGAATAAGTCTCTCAACATTAGGTCCCATTTTGGAGCCAGTTTC 1172

Db 3292 CAAGTGGAGTGAACCTTCGGAATAAGTCTCTCAACATTAGGTCCCATTTTGGAGCCAGTTTC 3351

QY 1173 TGACAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 1232

Db 3352 TGACAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA 3411

QY 1233 AGATGATGAATTAAGCCGCGAGGCACCTATTGAGGCGACCTTTCCAGCAGTTCCAGAAGCA 1292

Db 3412 AGATGATGAATTAAGCCGCGAGGCACCTATTGAGGCGACCTTTCCAGCAGTTCCAGAAGCA 3471

QY 1293 GAACGATGTACATAGGGCCCTTCAAGAGGGAAATTGAATAAAGAACTGTATATCATGAG 1352

Db 3472 GAACGATGTACATAGGGCCCTTCAAGAGGGAAATTGAATAAAGAACTGTATATCATGAG 3531

QY 1353 TACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACT 1412

Db 3532 TACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACT 3591

QY 1413 CTACCAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAGAGCCAGAAATGTCACTCCGCTTCT 1472

Db 3592 CTACCAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAGAGCCAGAAATGTCACTCCGCTTCT 3651

QY 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAATTTGAACCTGCACCTCCGCTGA 1532

Db 3652 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAATTTGAACCTGCACCTCCGCTGA 3711

QY 1533 CTGGCAGAGAAAAATAGATGAGACCTTTGAAAAGACTCCAGGAACCTTCAAGAGGCCACGGA 1592

Db 3712 CTGGCAGAGAAAAATAGATGAGACCTTTGAAAAGACTCCAGGAACCTTCAAGAGGCCACGGA 3771

QY 1593 TGAGCTGGACCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 1652

Db 3772 TGAGCTGGACCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 3831

QY 1653 CGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 1712

Db 3832 CGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 3891

QY 1713 AATTGGCCTCTGAAAAGAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCAC 1772

Db 3892 AATTGGCCTCTGAAAAGAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCAC 3951

QY 1773 TTTGGGCATTCAGCTCTCACCGTATACCTCAGCACTCTGGAAGACCTTGAACACCATG 1832

Db 3952 TTTGGGCATTCAGCTCTCACCGTATACCTCAGCACTCTGGAAGACCTTGAACACCATG 4011

QY 1833 GAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCCAGTCCAGTGCATGAAGCCACAGGGA 1892

Db 4012 GAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCCAGTCCAGTGCATGAAGCCACAGGGA 4071

QY 1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTGTCTCCAGGTCCCTGGGAGAGAGC 1952

Db 4072 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTGTCTCCAGGTCCCTGGGAGAGAGC 4131

QY 1953 CATCTCGCCAAACAAAGTGCCCTACTATATCAACACAGAGACTCAAACA 2001

Db 4132 CATCTCGCCAAACAAAGTGCCCTACTATATCAACACAGAGACTCAAACA 4180

RESULT 5

AX538624

LOCUS

DEFINITION

AX538624

ACCESSION

VERSION

AX538624.1 GI:25271175

Sequence 44 from Patent WO0229056.

AX538624

11443 bp

DNA

linear

PAT 23-NOV-2002

| | | | |
|------------|--|---|--------|
| QY | 1053 | ATCCTGGAAGGTTCCGATGATGCAGTCCTCTTTACAAAGACGTTTGGATAACATGAACCTT | 1112 |
| Db | 8500 | ATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACCTT | 8559 |
| QY | 1113 | CAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC | 1172 |
| Db | 8560 | CAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC | 8619 |
| QY | 1173 | TGACCAAGTGAAGCGCTGTCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA | 1232 |
| Db | 8620 | TGACCAAGTGAAGCGCTGTCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA | 8679 |
| QY | 1233 | AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGGAGTTCAGAAGCA | 1292 |
| Db | 8680 | AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGGAGTTCAGAAGCA | 8739 |
| QY | 1293 | GAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAACTAAAGAACCTGTAAATCATGAG | 1352 |
| Db | 8740 | GAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAACTAAAGAACCTGTAAATCATGAG | 8799 |
| QY | 1353 | TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACT | 1412 |
| Db | 8800 | TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACT | 8859 |
| QY | 1413 | CTACCAAGAGCCCGAGAGAGCTGCCTCTCTGAGGAGAGAGCCAGAAATGTCACCTCGGCTTCT | 1472 |
| Db | 8860 | CTACCAAGAGCCCGAGAGAGCTGCCTCTCTGAGGAGAGAGCCAGAAATGTCACCTCGGCTTCT | 8919 |
| QY | 1473 | ACGAAGCAGGCTGAGGAGGTCAATACTAGTGGGGAATAATTGAACCTGCACCTCCGCTGA | 1532 |
| Db | 8920 | ACGAAGCAGGCTGAGGAGGTCAATACTAGTGGGGAATAATTGAACCTGCACCTCCGCTGA | 8979 |
| QY | 1533 | CTGGCAGAGAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGA | 1592 |
| Db | 8980 | CTGGCAGAGAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGA | 9039 |
| QY | 1593 | TGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG | 1652 |
| Db | 9040 | TGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG | 9099 |
| QY | 1653 | CGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA | 1712 |
| Db | 9100 | CGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA | 9159 |
| QY | 1713 | AATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC | 1772 |
| Db | 9160 | AATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC | 9219 |
| QY | 1773 | TTTGGGCATTACGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATG | 1832 |
| Db | 9220 | TTTGGGCATTACGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATG | 9279 |
| QY | 1833 | GAAGCTTCTGAGGTGGCCCGTTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCCACAGGGA | 1892 |
| Db | 9280 | GAAGCTTCTGAGGTGGCCCGTTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCCACAGGGA | 9339 |
| QY | 1893 | CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCCTCGGAGAGAGC | 1952 |
| Db | 9340 | CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCCTCGGAGAGAGC | 9399 |
| QY | 1953 | CATCTCGCCAAACAAGTGCCCTACTATATCAACACCAGAGACTCAAACA | 2001 |
| Db | 9400 | CATCTCGCCAAACAAGTGCCCTACTATATCAACACCAGAGACTCAAACA | 9448 |
| RESULT 7 | | | |
| HSDMDR | | | |
| LOCUS | HSDMDR | 12446 bp | linear |
| DEFINITION | Human mRNA for dystrophin. | | |
| ACCESSION | X14298 | | |
| VERSION | X14298.1 | GI:30845 | |
| KEYWORDS | Dmd gene; Duchenne muscular dystrophy; dystrophin. | | |

| | |
|-----------|--|
| SOURCE | Homo sapiens (human) |
| ORGANISM | Homo sapiens |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. |
| AUTHORS | 1 (bases 1 to 12446) Rosenthal,A., Speer,A., Billwitz,H., Cross,G.S., Forrest,S.M. and Davies,K.E. |
| TITLE | Two human cDNA molecules coding for the Duchenne muscular dystrophy (DMD) locus are highly homologous |
| JOURNAL | Nucleic Acids Res. 17 (13), 5391 (1989) |
| MEDLINE | 89345106 |
| PUBMED | 2668885 |
| REFERENCE | 2 (bases 1 to 12446) Rosenthal,A. |
| AUTHORS | Direct Submission |
| TITLE | Submitted (09-FEB-1989) Rosenthal A., Akademie der Wissenschaften |
| JOURNAL | der DDR, Zentralinstitut fuer Molekularbiologie, Robert-Roessle Str.10, 1115 Berlin Buch, DDR |
| COMMENT | see also M18533 and M20250 for Dmd seqs.; discrepancies compared to M18533 cDNA were located at x14298 pos. 496, 1772, 1965, 2449, 3687, 4229, 4504, 5075, 5332, 5630 and 7194. |
| FEATURES | Location/Qualifiers 1..12446 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /chromosome="X chromosomal, xp21." /tissue_type="muscle" /dev_stage="fetal and adult." 99..11156 /note="dystrophin (AA 1 - 3685)" /codon_start=1 /protein_id="CAA32479.1" /db_xref="GI:30846" /db_xref="SWISS-PROT:P11532" /translation="MLWVEEVEDCYEREDVQKTFKWNVAQFSKFGKHQIENLFSDL QDGRLLDLLLEGTLQKLPKEKSTRVHALNNVNNKALRVLQNNVNDLVNIGSTDIVDG NKKLTGLIWNIIILHWQKNVMKNIMAGLOPTNSEKILLSWVRQSTRNPQVNVINFT TSWSDGLALNALJHSHRPDLFDWNSVVCOSATQRLHAFNARIYQLGIEKLDLPEDV DTTPDKKSIILMYTISLPQVLPOVSIEAIQEVEMLPFPKVTKEEHFQLHQMHYSQ QITVSLAQGYERTSPKPRFKSYAYTQAAVYTTSDPTSPFPSPHQLEAPEDKSFSSSL MESEVNLDRYQTALBEVLSWLLSAEDTLQAQGEISNDVEVVKDQFHHHEGYMDLTAH QGRVGNILQLSGLITGKLTSEDETEVOQMNLNLSRWECLRVASMEKQSNLHRVLM DLQNKKLKELNDWLTKEERTRMEEEEPLGPDLEDLKRQVQKHVKVLEDEEQEVVRVN SLTHMVVWVDESSGDHATAALEQLKVLGDRWANI CRWTEDRWVLLQDILLKQRLTE EQCLFSAWLSEKEDAVNKIHTTGFKDQNMELSSLOKLAVLKADLEKKQSMGKLYSIK QDLLSTLKNKSVTQKTEAWLDFARCDNLVQLEKLEKTAQISOAVTTTPSLTQTIVM ETVTVTTRREQILVHAQEELPPPPQKQKQITVDSEIRKRLDVIDTELHWSITRSEA VLQSPFAIFRKEGNEFDLKEKVNAIEREKAERFKRLQDASRSGQALVEQMVNEGVNA DSIKQASEQLNSRWIEFCQLLSERLNWLEYQNNIIAFYNQLQQLEQMTTAAENWLKIQ PTPPSEPTAIKSQLKICKDEVNRLSGLQPIERLKIQSIALKEKGQGMFLDADFVAF TNHFQKQVFSQVQAREKELQITFDLTPPMRYQETMSAIRTWVQQSETKLSIPQLSVTDY EIMEQRLGELQALQSSLEQQSGLYLSTTVKEMSKKAPSEISRKYQSEFEIEIEGRWK KLSQLVEHCQKLEQMNKLRKIQNHQITLKKWMAEVDVFLKEWPALGDSIELKKQL KQCRLLVSDIQTIQPSLNSVNEGQKIKNEAEPEFASRLTELELTELKELNTOWDHMCQQVY ARKEALKGGLEKTVSLQKDLSEMHWMTOAEEYLERDFFEYKTPELQKAFHEMKRAK EEAQQKBAKVLLTESVNSVIAQAPPVAQEAALKKELETLTTNYQWLCTRLNGKCKTLE EVMACWHELLLSYLEKANWLNVEVFKLTETENIPQGAEEISEVLDLSLENLMRHSNDNP NQIRILAQTLTDGGVMDELINEELETFSNRWRLEHAEVRRQKLEQSIQSAQSETNS LHLIQESLTFIDKQLAAVADKVDAQMPQEAQIKTQSDLTSHETISLEEMKHNQCKEA AQRVLSQIDVAQKQLQDSVMKFRLLFOKPAFNEQRLQESKMWLDEVKMLHPALETKSVE QEVVQSQNLHCVNLYKSLSEVKSEVEMVIKTGRQIVQKKQTPENPKELDERVTLKLHY NELGAKVTERKQKLEKCLKLSRKMRKEMNVLTEWLATMDMELTKRSAVEGMPNSLDSH VAWGKATQKEIEKQKVHLKSITEVGEALKTVLGKKELTLEDKLSLLNSNWIATVSRAE EWLNLILEYQKHMETFQNVNDHITKWI IQADTLDESEKKKPPQKEDVLKRLKAELND IRPKVDSRDRQQAANLMAHNDHCRKLVEPQISELNHRFAAISHRITKGASIPLKELE QFNSDIQKLEPLEAEIQQGVNLKEEDFNKDMNEDNEGVKELLQRGDNLQQRITDER KSEETIKQQLLQTKHNALKDLRSQRRKKALEISHOWYQKRQADDLLKCLDDIEKKL ASLPEPRDERKIKEIDRELQKKEELNAVRRQAEGLSGDGAAMAVEPTQIOLSKRWE IESKFAQFRRRLNFAQIHRTVREETMMVTMEDMPLEISYVPSTYLTETITVHSQALLEVEQ LLNAPDLCAKDFEDLFKQEEELKNIKDSLQSSSGRIDIHSKKTAALQSATPVERVKL QEALSQLDFQWEKVNKMYKDRQGRFDRSVEKWRFRHYDIKIFNQWLTEAEQFLRKTQI PENWEHAKYKWLKELQDGGIGQRQTVVVRTNLNATGEEIIQOSSKTDASILQEKLGSLNL |
| CDS | |

BASE COUNT 4135 a 2524 c 2876 g 2911 t
ORIGIN
Query Match 59.1%; Score 1182.6; DB 9; Length 12446;
Best Local Similarity 99.7%; Pred. No. 7.8e-286;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAAGTTTCTTGCTG 872
DB 8150 GGAAGAACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAAGTTTCTTGCTG 8209
QY 873 GCTTACAGAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGGCT 932
DB 8210 GCTTACAGAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGGCT 8269
QY 933 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAAATGGCAAGACCTCCAAGTGA 992
DB 8270 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAAATGGCAAGACCTCCAAGTGA 8329
QY 993 AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAGCCCAAAAATCTCTGAG 1052
DB 8330 AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAGCCCAAAAATCTCTGAG 8389
QY 1053 ATCCCTGGAAGGTTCCGATGATGCAGTCTCTCAACATTAGGTCCCATTTGGAATACATGAAC 1112
DB 8390 ATCCCTGGAAGGTTCCGATGATGCAGTCTCTCAACATTAGGTCCCATTTGGAATACATGAAC 8449
QY 1113 CAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTTC 1172
DB 8450 CAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTTC 8509
QY 1173 TGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGGCTACAGCTGAA 1232
DB 8510 TGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGGCTACAGCTGAA 8569
QY 1233 AGATGATGAATTAAGCCGCGAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAAGCA 1292
DB 8570 AGATGATGAATTAAGCCGCGAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAAGCA 8629
QY 1293 GAACGATGTACATAGGCGCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTATCATGAG 1352
DB 8630 GAACGATGTACATAGGCGCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTATCATGAG 8689
QY 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGACGAGCCCTTTGGAAGGACTAGAGAACT 1412
DB 8690 TACTCTTGAGACTGTACGAATATTTCTGACAGACGAGCCCTTTGGAAGGACTAGAGAACT 8749

QY 1413 CTACCAGGAGCCAGAGAGTGCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT 1472
DB 8750 CTACCAGGAGCCAGAGAGTGCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT 8809
QY 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTCGCTGA 1532
DB 8810 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTCGCTGA 8869
QY 1533 CTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACCGA 1592
DB 8870 CTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACCGA 8929
QY 1593 TGAGCTGACCTCAAGCTGCGGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGG 1652
DB 8930 TGAGCTGACCTCAAGCTGCGGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGG 8989
QY 1653 CGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 1712
DB 8990 CGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 9049
QY 1713 AATTGGCGCTCTGAAAGAGAACGTGAGCCACCGTCAATGACCTTGTCTCGCAGCTTACCAC 1772
DB 9050 AATTGGCGCTCTGAAAGAGAACGTGAGCCACCGTCAATGACCTTGTCTCGCAGCTTACCAC 9109
QY 1773 TTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACAGATG 1832
DB 9110 TTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACAGATG 9169
QY 1833 GAAGCTTCTGCAGGTGGCGCTCGAGACCGAGTCAGGCAGCTGCATGAAGCCACACAGGA 1892
DB 9170 GAAGCTTCTGCAGGTGGCGCTCGAGACCGAGTCAGGCAGCTGCATGAAGCCACACAGGA 9229
QY 1893 CTTTGGTCCAGCATCTCAGCACTTTTCTTCCACGTCTGTCCAGGTCCTTGGGAGAGAGC 1952
DB 9230 CTTTGGTCCAGCATCTCAGCACTTTTCTTCCACGTCTGTCCAGGTCCTTGGGAGAGAGC 9289
QY 1953 CATCTGCCAAACAAAGTGCCCTACTATATCAACACCGAGACTCAAACA 2001
DB 9290 CATCTGCCAAACAAAGTGCCCTACTATATCAACACCGAGACTCAAACA 9338
RESULT 8
AX409637
LOCUS
DEFINITION Sequence 2284 from Patent WO0229103.
ACCESSION AX409637
VERSION AX409637.1 GI:21442342
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS
TITLE
JOURNAL
Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.
Gene expression profiles in liver cancer
Patent: WO 0229103-A 2284 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
Location/Qualifiers
1. .13957
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. M18533"
BASE COUNT 4602 a 2781 c 3122 g 3452 t
ORIGIN
Query Match 59.1%; Score 1182.6; DB 6; Length 13957;
Best Local Similarity 99.7%; Pred. No. 7.9e-286;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAAGTTTCTTGCTG 872
DB 8260 GGAAGAACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAAGTTTCTTGCTG 8319

1413 CTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT 1472
|||||
8860 CTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT 8919
|||||
1473 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTGCACTCCGCTGA 1532
|||||
8920 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTGCACTCCGCTGA 8979
|||||
1533 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGA 1592
|||||
8980 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGA 9039
|||||
1593 TGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 1652
|||||
9040 TGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 9099
|||||
1653 CGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGAGAAAAGTCAAGGCACTTCGAGGAGA 1712
|||||
9100 CGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGAGAAAAGTCAAGGCACTTCGAGGAGA 9159
|||||
1713 AATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCAC 1772
|||||
9160 AATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCAC 9219
|||||
1773 TTTGGGCATTGAGCTCTCACCCTATAACCTCAGCACTCTGGAAGACCTTGAACACCCAGATG 1832
|||||
9220 TTTGGGCATTGAGCTCTCACCCTATAACCTCAGCACTCTGGAAGACCTTGAACACCCAGATG 9279
|||||
1833 GAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGACGCTGCATGAAGCCACAGGGA 1892
|||||
9280 GAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGACGCTGCATGAAGCCACAGGGA 9339
|||||
1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTGTGTCAGGGTCCCTGGGAGAGAGC 1952
|||||
9340 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTGTGTCAGGGTCCCTGGGAGAGAGC 9399
|||||
1953 CATCTCGCCAAACAAAGTGCCCTACTATATCAACACGAGACTCAAACA 2001
|||||
9400 CATCTCGCCAAACAAAGTGCCCTACTATATCAACACGAGACTCAAACA 9448
|||||

RESULT 10
HUMDYS
LOCUS HUMDYS 13957 bp mRNA linear PRI 25-MAY-2000
DEFINITION Homo sapiens dystrophin (DMD) mRNA, complete cds.
ACCESSION M18533 M17154 M18026 M20250
VERSION M18533.1 GI:181856
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1699)
Koenig,M., Hoffman,E.P., Bertelson,C.J., Monaco,A.P., Feener,C. and
Kunkel,L.M.
Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and
preliminary genomic organization of the DMD gene in normal and
affected individuals
Cell 50 (3), 509-517 (1987)
JOURNAL MEDLINE 87273512
PUBMED 3607877
REFERENCE 2 (bases 1678 to 3830)
Hoffman,E.P., Monaco,A.P., Feener,C.C. and Kunkel,L.M.
AUTHORS Conservation of the Duchenne muscular dystrophy gene in mice and
humans
JOURNAL Science 238 (4825), 347-350 (1987)
MEDLINE 88018015
PUBMED 3659917
REFERENCE 3 (bases 1 to 13957)
Koenig,M., Monaco,A.P. and Kunkel,L.M.
AUTHORS The complete sequence of dystrophin predicts a rod-shaped
TITLE cytoskeletal protein
JOURNAL Cell 53 (2), 219-226 (1988)

MEDLINE 88194521
PUBMED 3282674
COMMENT On May 25, 2000 this sequence version replaced gi:340693.
Draft entry and computer-readable sequence kindly provided by
M.Koenig, 01-APR-1988 The severity of muscular dystrophy is
determined by the size of the deleted DNA segment. Deletions found
in different patients were from positions 302-2200, 473-1168,
1691-1810, and 1169-3011.
Location/Qualifiers
1. .13957
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp21.3-p21.1"
/tissue_type="muscle"
/dev_stage="fetus"
/note="G00-119-850"
1. .13957
/gene="DMD"
209. .11266
/gene="DMD"
/codon_start=1
/product="dystrophin"
/protein_id="AAA53189.1"
/db_xref="GI:181857"
/translation="MLWEEVEDCYEREDVQKKTFTKWNAQFSKFKQKHENLFSDL
ODGRRLLDLEGLTGQKLPEKGGSTRVHALNVNKKALRVLONNVNVLNIGSTDIVDG
NHKLTGLIWNILHWQKNVMKNIMAGLQQTNSEKILLSVWRQSTRNYPQVNVINFT
TSWSDGLALNALHSHRPDLFDWNSVVCQOSATORLEHAFNARIYQIGIEKLLDPEDV
DITYDPKKSILMYITSLFQVLPQQVSTEAIOEVEMLPRPKVTKEHFQLHQMHSQ
QITVSLAQGYERTSSPKPRFSYAYTAAAVTSDPTRSPFSPQHLPAPEKDSFGSSL
MESEVNLDRYQFALIEVLSLLSAEDTLQAQGEISNDVEVVKDQFTHHEGYMMDLTAH
QGRVGNILQGSKLIGTKLSEDEETEVEQMNLLNSRWECLRVASMEKQSNLHRVLM
DLQNKLLKELNDWLTKEERTRKMEEPGLPDLEDLKRQVQKHVQLQEDLEQEVQVRV
SLTHMVVVDESSGDHATAALEEQKVLGDRWANICRWTEDRWVLLQDILLKWQRLTE
EQCLFSAWLSEKEDAVNKHITTGFKDQNEMLSSLQKLAVLKADLEKKQSMGKLYSLK
QDLLSTLKNKSVTKTEAWLDFARNLQVLEKSTAQISQAVTTTOPSLTQTVM
ETVTVTTRQILVKHAQEELPPPPQPKKQITVDSEIRKRLDVIDITELHSWITRSEA
VLQSPFAIFRKEGNFSDLKEVNAIEREKAERKFLQDASRSAQALVEQMVNIEGVNA
DSIKQASEQLNSRWIEFCOLLSERLNWLEYQNNIIAFYNLQQLEQMTTFAENWLKIQ
PTTSEPPTAKSQLKICKDEVNRLSGLQPIERLKIQSIALKEKGQGPFLDADFVAF
TNHFQVFSVDQAREKELQITFDTLPPMYQETMSAIRTWVQQSETKLSIPQLSVTDY
ETMEQRLGELQSSSQEQSGLYLSTTVKMSKAPSEISRKYQSEFEEIEGRWK
KLSSQLVEHCQKLEEQMNKLRKIONHITLKKWMAEVDVFLKEEWPALGDSEILKKQL
KQCRLLVSDIQTTPSLNSVNEGGQIKNEAPEFASRLTELTTELKELNTQDHWMCQVY
ARKEALKGGLKTVSLQKDLSEHWMVTQAEVEYLERDFEYKTPDELQKAVEEMKRAK
EEAQKKEAKVKLLTESVNSVIAQAPPVQAQALKALETLTITNYQWLCTRLNGKCKTLE
EWWACWHELLSYLEKANKWLNEVEFKLKTENIPGGAEEISEVLDLSLENLMRHSEDNP
NQIRILAQTLTDDGVNDELINEELETFNSRWRELHEEAVRRQKLEQSIQSAQETESK
LHLIOESLTFIDKQLAAYIADKVDAAQMPQEAQIQSDLSLTSHEISLEEMKKNQKGEA
AQRVLSQIDVDAQKQLQDVSMKFRFLFQKPFANFELRQESKMLIDEVKMLPALETKSVE
QEVVQSQLNHCVNLYKSLSEVKSEVEMVIKTGRIQIVQKKQNTENKELDERVTALKLHY
NELGAKVTERKQOQLEKCLKLSRKMRKEMNVLTEWLAATDMELTKRSAVEGMPNLDSE
VANGKATQKEIEKQKVLKSI TEVGEALKTVLGKKTVELDKLSLLNSNWIATVSRAE
EWNLLLELYQKHMETFDONVDHITKWI IQADTLLDSESEKKKPPQKQEDVLKRLKAELND
IRPKVDSTRDQAANLMNRGDHCRKLVEPOISELNRFAAI SHRIKTKASTPLKLELE
QFNSDIQKLLLEAEIQOQVNLKEEDFNKDMNEDNEGTVKELLQRGDNLQQRITDER
KREEIKIQQLLQTKHNAKDLRSQRRKALEISHQWYQYKQADDDLKCLDDIEKKL
ASLPEPRDERKIKEIDRELQKKEELNARRQAEGLSEGDGAMA VEPTQIQLSKRWRE
TESKFAQRRRLNFAQIHTVREETMMVMTEDMPLEISYVPSTYLTEITHVSOALLVEVEQ
LLNAPDLCAKDFEDLFKOEELKNIKDSLOQSSGRIDI IHSKTAALQASATPVERVKL
QEALSQLDFQWEKVNMYKDRQGRFDRSVEKWRFRHYDIKIFNQWLTEAEQFLRKTQI
PENWEHAKYKWLKELQDGIGQRQTVRTLNTATGEI IQSSSKTDASILQEKLSLNL
RWQEVCKQLSDRKRLEEQKNILSEFQDLNLFVLEADNIASIPLEPGKEQQLKVS
KLEQVKLLVEELPLRQGLKQNETGGPVLVSAPISPEEQKLENKLNQNLQWIKVS
RALPEKQGEIEAQIKOLGQLEKKLEDLEEQNLHLLWLSPIRNQLEIYNQPNQEGPFD
VQETI AVQAKQPDVVEILSKQHLKYEKPATQPKRKLDELSEWKA VNRLLQELRA
KQPD LAPGLTTIGASPTQTVTLTPQVPVTKETAISKLEMPSSLMLEVPALADFNRAWT
ELTDWLSLLDQVIKSORVMVGDLIEDINEMI IKQATMODLEQRRPQLEELITAAQNLIK
NKTSNQEARITITDRIERIQNQWDEVQEHQNRRLQNLNEMLKDSTQWLEAKEEEAEQVL
GQARAKLESWKEGPTYVDAIQKKITETKOLAKDLRQWQTNVDVANDLALKLRDYSAD
DTRKVEHMITENINASWRSIHKRVSREAALEETHRLLOQFPDLLEKFLAWLTEAETTA

| | | | |
|----------------------------|--------|--|---------------------|
| | | NVLQDATRKRLLLEDSKGVKELMKQWDLQGEIEAHTDVVHNLNDSQKILRSLEGS | |
| | | DAVLLQRLDNNFKNSELRKSLNIRSHLEASSDQWKRHLHLSQELLVWLQKXDDL | |
| | | SROAPIGGDFPAVKQNDVHRAFKRELKTKPEVIMSTLETVRIFLTEQPLEGLEKLYQ | |
| | | EPRELPPERBAQNVTRLLRKQAEVNTWEKLNHLSADWOKRIDETLERLQELQRAVD | |
| | | ETDLKLRQAEVIKGSWPVGDLISLDQHLKVKALRGELAPLKENVSHVNDLARQL | |
| | | TTIGIQSPNLSTLEDLNRWKLQVAVEDRVQLHEARDFFPASQHFSTLSVQGP | |
| | | WERAISPNKVPYIYNHETQTTDWDHPKMTLYQSLADLNNVRFSAYRTAMKLRLLQKA | |
| | | LCLDLLSLSACDALDQHLNKQNDQMDILOIINCLTTIYDRLEQEHNNLVNPLCVD | |
| | | MCLNLLNVYDTRGRIRVLSFKTIGIISLCKAHLEDKRYLFKQVASSTGFCQDRRL | |
| | | GLLLHDSIOIPRQIJEVASFGSNIEPSVRSCFOFANNKPEIEAALFLDWMRLPQSM | |
| | | WLPVLHVRVAAETAKHQAKNICKEPIIGFPRYSLKHFNVDICQSCFFSGRVAKGH | |
| | | KMHPMVEYCTPTTSGEDVRDEAKVLKNKFRKRYFAKHPRMGYLPVQTVLEGDNMET | |
| | | PVTLINFPVDSAPASSPQLSHDDTHSRIEHYASRLAEMENSGSYLNDISIPNESID | |
| | | DEHLLIQHYCQSLNQDSPLSQSPSPAQILISLESEERGELERILADLEEENRNLQAEY | |
| | | DRLKQHEHKGLSPSPPEMPTSPQSPRDAELIAEAKLLRQHKGRLEARMQIILEDH | |
| | | NKQLESQHLRLQLLEQPOAEAKVNGTTVSPSTSLQRSDDSOPLLRVVGSQTSDSM | |
| | | GEEDLLSPPDQSTGLEEVMEOQLNNSFPSSRGRNTPGKPMREDTM" | |
| BASE COUNT | 4602 a | 2781 c | 3122 g 3452 t |
| ORIGIN | | | |
| Query Match | 59.1%; | Score 1182.6; | DB 9; Length 13957; |
| Best Local Similarity | 99.7%; | Pred. No. 7.9e-286; | |
| Matches 1185; Conservative | 0; | Mismatches 4; | Indels 0; Gaps 0; |
| QY | 813 | GGTACCTACTCATAGATTACTGCAACAGTTTCCCCCTGGACCTGGAAGATTTTCTTGCCTG | 872 |
| Db | 8260 | GGAAGAACTCATAGATTACTGCAACAGTTTCCCCCTGGACCTGGAAGATTTTCTTGCCTG | 8319 |
| QY | 873 | GCTTACAGAACTGAAACAACTGCCAATGTCTTACAGGATGCTACCCGTAAAGAAAGGCT | 932 |
| Db | 8320 | GCTTACAGAACTGAAACAACTGCCAATGTCTTACAGGATGCTACCCGTAAAGAAAGGCT | 8379 |
| QY | 933 | CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAGACCTCCAAGGTGA | 992 |
| Db | 8380 | CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAGACCTCCAAGGTGA | 8439 |
| QY | 993 | AATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAACAGCCCAAAAAATCCTGAG | 1052 |
| Db | 8440 | AATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAACAGCCCAAAAAATCCTGAG | 8499 |
| QY | 1053 | ATCCCTGGAAGGTTCCGATGATGACCTTGTACAAAGACGTTTGGATAACATGAACCTT | 1112 |
| Db | 8500 | ATCCCTGGAAGGTTCCGATGATGACCTTGTACAAAGACGTTTGGATAACATGAACCTT | 8559 |
| QY | 1113 | CAAGTGGAGTGAACTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC | 1172 |
| Db | 8560 | CAAGTGGAGTGAACTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC | 8619 |
| QY | 1173 | TGACCAAGTGAAGCGTCTGCACCTTTCTCTGAGGAACTTCTGCTGCTGCTACAGCTGAA | 1232 |
| Db | 8620 | TGACCAAGTGAAGCGTCTGCACCTTTCTCTGAGGAACTTCTGCTGCTGCTACAGCTGAA | 8679 |
| QY | 1233 | AGATGATGAATTAAGCCGCGAGGCACCTATTGGAGCGGACCTTTCCAGAGCTTCAGAAAGCA | 1292 |
| Db | 8680 | AGATGATGAATTAAGCCGCGAGGCACCTATTGGAGCGGACCTTTCCAGAGCTTCAGAAAGCA | 8739 |
| QY | 1293 | GAAAGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTATCATGAG | 1352 |
| Db | 8740 | GAAAGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTATCATGAG | 8799 |
| QY | 1353 | TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT | 1412 |
| Db | 8800 | TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT | 8859 |
| QY | 1413 | CTACCAAGAGCCAGAGAGCTGCTCCTGAGGAGAGAGCCAGAAATGTCTACTCGGCTTCT | 1472 |
| Db | 8860 | CTACCAAGAGCCAGAGAGCTGCTCCTGAGGAGAGAGCCAGAAATGTCTACTCGGCTTCT | 8919 |
| QY | 1473 | ACGAAAGCAGCTGAGGAGGTCAATACTAGTGGGAAAAAATGAACCTGCTCCGCTGA | 1532 |
| Db | 8920 | ACGAAAGCAGCTGAGGAGGTCAATACTAGTGGGAAAAAATGAACCTGCTCCGCTGA | 8979 |
| QY | 1533 | CTGGCAGAGAAAAATAGATGAGACCTTGTAAAGACTCCAGGAACTTCAAGAGGCCACGGA | 1592 |

| | | | |
|----------------------------|--------|--|----------------------------|
| Db | 8980 | CTGGCAGAGAAAAATAGATGAGACCTTGTAAAGACTCCAGGAACCTTCAAGAGGCCACGGA | 9039 |
| QY | 1593 | TGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG | 1652 |
| Db | 9040 | TGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG | 9099 |
| QY | 1653 | CGATCTCTCATTTGACTTCTTCCAAGATCACTCGAGAAAGTCAAGGCACTTCGAGGAGA | 1712 |
| Db | 9100 | CGATCTCTCATTTGACTTCTTCCAAGATCACTCGAGAAAGTCAAGGCACTTCGAGGAGA | 9159 |
| QY | 1713 | AATTGGCCTCTGAAAGAGAACTGTAGCCACGTCACGTCACCTTGCTGCGCAGCTTACCAC | 1772 |
| Db | 9160 | AATTGGCCTCTGAAAGAGAACTGTAGCCACGTCACGTCACCTTGCTGCGCAGCTTACCAC | 9219 |
| QY | 1773 | TTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACAGATG | 1832 |
| Db | 9220 | TTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACAGATG | 9279 |
| QY | 1833 | GAGCTTCTGAGGTGGCCCTCGAGGACCGAGTCAAGGAGCTGCATGAAGCCACACAGGGA | 1892 |
| Db | 9280 | GAGCTTCTGAGGTGGCCCTCGAGGACCGAGTCAAGGAGCTGCATGAAGCCACACAGGGA | 9339 |
| QY | 1893 | CTTGGTCCAGCATCTCAGCACTTCTTTCCACGCTGTCTCCAGGGTCCCTGGGAGAGAGC | 1952 |
| Db | 9340 | CTTGGTCCAGCATCTCAGCACTTCTTTCCACGCTGTCTCCAGGGTCCCTGGGAGAGAGC | 9399 |
| QY | 1953 | CATCTGCCAAACAAAGTGCCTTACTATATCAACCACGAGACTCAAAACA | 2001 |
| Db | 9400 | CATCTGCCAAACAAAGTGCCTTACTATATCAACCACGAGACTCAAAACA | 9448 |
| RESULT 11 | | | |
| AR220819 | | 13977 bp | DNA linear PAT 26-SEP-2002 |
| LOCUS | | Sequence 60 from patent US 6426186. | |
| DEFINITION | | AR220819 | |
| ACCESSION | | AR220819 | |
| VERSION | | AR220819.1 | GI:23327696 |
| KEYWORDS | | Unknown. | |
| SOURCE | | Unknown. | |
| ORGANISM | | Unclassified. | |
| REFERENCE | | 1 (bases 1 to 13977) | |
| AUTHORS | | Jones, K.A., Volkmuth, W. and Walker, M.G. | |
| TITLE | | Bone remodeling genes | |
| JOURNAL | | Patent: US 6426186-A 60 30-JUL-2002; | |
| FEATURES | | Location/Qualifiers | |
| source | | 1. 13977 | |
| BASE COUNT | 4596 a | 2765 c | 3120 g 3453 t |
| ORIGIN | | /organism="unknown" | 43 others |
| Query Match | 59.1%; | Score 1182.6; | DB 6; Length 13977; |
| Best Local Similarity | 99.7%; | Pred. No. 7.9e-286; | |
| Matches 1185; Conservative | 0; | Mismatches 4; | Indels 0; Gaps 0; |
| QY | 813 | GGTACCTACTCATAGATTACTGCAACAGTTTCCCCCTGGACCTGGAAGATTTTCTTGCCTG | 872 |
| Db | 8260 | GGAAGAACTCATAGATTACTGCAACAGTTTCCCCCTGGACCTGGAAGATTTTCTTGCCTG | 8319 |
| QY | 873 | GCTTACAGAACTGAAACAACTGCCAATGTCTTACAGGATGCTACCCGTAAAGAAAGGCT | 932 |
| Db | 8320 | GCTTACAGAACTGAAACAACTGCCAATGTCTTACAGGATGCTACCCGTAAAGAAAGGCT | 8379 |
| QY | 933 | CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAGACCTCCAAGGTGA | 992 |
| Db | 8380 | CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAGACCTCCAAGGTGA | 8439 |
| QY | 993 | AATTGAAGCTCACACAGATGTTTATCAACAACCTGGAATGAAACAGCCCAAAAAATCCTGAG | 1052 |
| Db | 8440 | AATTGAAGCTCACACAGATGTTTATCAACAACCTGGAATGAAACAGCCCAAAAAATCCTGAG | 8499 |
| QY | 1053 | ATCCCTGGAAGGTTCCGATGATGACCTTGTACAAAGACGTTTGGATAACATGAACCTT | 1112 |

```

8500 ATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAAC TT 8559
1113 CAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTTGGAGCCAGTTC 1172
8560 CAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTTGGAGCCAGTTC 8619
1173 TGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAA 1232
8620 TGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAA 8679
1233 AGATGATGAATTAAGCCGGCAGGACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAAAGCA 1292
8680 AGATGATGAATTAAGCCGGCAGGACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAAAGCA 8739
1293 GAACGATGTACATAGGCGCTTCAAGAGGGGAATTGAAACTAAAGAACCTGTATCATGTAG 1352
8740 GAACGATGTACATAGGCGCTTCAAGAGGGGAATTGAAACTAAAGAACCTGTATCATGTAG 8799
1353 TACTCTTGAGACTGTACGAATATTCTTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACT 1412
8800 TACTCTTGAGACTGTACGAATATTCTTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACT 8859
1413 CTACCAGAGCCAGAGAGCTGCTCCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT 1472
8860 CTACCAGAGCCAGAGAGCTGCTCCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT 8919
1473 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCACCTCCGCTGA 1532
8920 ACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCACCTCCGCTGA 8979
1533 CTGGCAGAGAAAAATAGATGAGAGCCCTTGAAAGACTCCAGSAACTTCAAGAGGCCACGGA 1592
8980 CTGGCAGAGAAAAATAGATGAGAGCCCTTGAAAGACTCCAGSAACTTCAAGAGGCCACGGA 9039
1593 TGAGCTGGACCTCAAGCTGCGGCCAAGCTGAGGTGATCAAGSGATCCTGGCAGCCCGTGGG 1652
9040 TGAGCTGGACCTCAAGCTGCGGCCAAGCTGAGGTGATCAAGSGATCCTGGCAGCCCGTGGG 9099
1653 CGATCTCCTCATTTGACTCTCTCCAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 1712
9100 CGATCTCCTCATTTGACTCTCTCCAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 9159
1713 AATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC 1772
9160 AATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC 9219
1773 TTTGGGCATTCAAGCTCTCACCGHATAACCTCAGCACTCTGGAAGACCTGAACACCGATG 1832
9220 TTTGGGCATTCAAGCTCTCACCGHATAACCTCAGCACTCTGGAAGACCTGAACACCGATG 9279
1833 GAAGCTTCTGCAGTGGCCGTTCGAGGACCGAGTCAGGCAGCTGATGAAGCCACAGGGA 1892
9280 GAAGCTTCTGCAGTGGCCGTTCGAGGACCGAGTCAGGCAGCTGATGAAGCCACAGGGA 9339
1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCCCTGGGAGAGAGC 1952
9340 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCCCTGGGAGAGAGC 9399
1953 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCCACGAGACTCAAAACA 2001
9400 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCCACGAGACTCAAAACA 9448
```

```

RESULT 12
AX538620
LOCUS AX538620 5339 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 40 from Patent WO0229056.
ACCESSION AX538620
VERSION AX538620.1 GI:25271166
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
```

```

artificial sequences.
REFERENCE
1 Chamberlain, J.S. and Harper, S.Q.
AUTHORS Mini-dystrophin nucleic acid and peptide sequences
TITLE Patent: WO 0229056-A 40 11-APR-2002;
JOURNAL THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
source
1.5339
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"
BASE COUNT 1638 a 1191 c 1187 g 1323 t
ORIGIN
Query Match 59.0%; Score 1180; DB 6; Length 5339;
Best Local Similarity 100.0%; Pred. No. 3.4e-285;
Matches 1180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 822 TCATAGATTACTGCAACAGTTCCCTCTGACCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGA 881
Db 1546 TCATAGATTACTGCAACAGTTCCCTCTGACCTGGAAAAAGTTTCTTGCTGGCTTACAGA 1605
QY 882 AGCTGAAACAACTGCCAATGTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAGA 941
Db 1606 AGCTGAAACAACTGCCAATGTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAGA 1665
QY 942 CTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 1001
Db 1666 CTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 1725
QY 1002 TCACACAGATGTTTATCAACAACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGA 1061
Db 1726 TCACACAGATGTTTATCAACAACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGA 1785
QY 1062 AGTTCCGATGATGAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTG 1121
Db 1786 AGTTCCGATGATGAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTG 1845
QY 1122 TGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTG 1181
Db 1846 TGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTG 1905
QY 1182 GAAGCGTCTGCACCTTTCTCTGACGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGA 1241
Db 1906 GAAGCGTCTGCACCTTTCTCTGACGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGA 1965
QY 1242 ATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTTCAGAAAGCAGAACGATGT 1301
Db 1966 ATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTTCAGAAAGCAGAACGATGT 2025
QY 1302 ACATAGGGCCTTCAAGAGGGAAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTTGA 1361
Db 2026 ACATAGGGCCTTCAAGAGGGAAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTTGA 2085
QY 1362 GACTGTACGAATATTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAAACTTACCAGGA 1421
Db 2086 GACTGTACGAATATTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAAACTTACCAGGA 2145
QY 1422 GCCCAGAGAGCTGCTCTCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCA 1481
Db 2146 GCCCAGAGAGCTGCTCTCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCA 2205
QY 1482 GGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAG 1541
Db 2206 GGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAG 2265
QY 1542 AAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGGA 1601
Db 2266 AAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGGA 2325
QY 1602 CCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGCGGATCTCCT 1661
```


1053 ATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTACAAAAGACGTTTGGATAACATGAACCTT 1112
1113 CAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTTGGAAAGCCAGTTC 1172
8626 CAAGTGGAGCGAACTTCGGAAAAAGTCTCTCAACATTAGGTCTCAGTTTGGAAAGCCAGTTC 8685
1173 TGACCATGTGAAGCGTCTGCACCTTCTCTGACGAACTTCTGGTGTGGTGTACAGCTGAA 1232
8686 TGACCATGTGAAGCGTCTGCACCTTCTCTGACGAACTTCTGGTGTGGTGTACAGCTGAA 8745
1233 AGATGATGAATTAAGCCGAGGACCTATTGGAGCGGACCTTCCAGCAGTTCAGAAAGCA 1292
8746 AGATGATGAATTAAGCCGAGGACCTATTGGAGCGGACCTTCCAGCAGTTCAGAAAGCA 8805
1293 GAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTGAATCATGAG 1352
8806 GAATGATGTACACAGGCGCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTGAATCATGAG 8865
1353 TACTCTTGAGACTGTACGAATATTCTGACAGAGCGGCTTTGGAAAGGACTAGAGAAACT 1412
8866 TACTCTTGAGACTGTACGAATATTCTGACAGAGCGGCTTTGGAAAGGACTAGAGAAACT 8925
1413 CTACAGAGCCAGAGAGCTGCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT 1472
8926 CTACAGAGCCAGAGAGCTGCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT 8985
1473 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGA 1532
8986 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGA 9045
1533 CTGGCAGAGAAAAATAGATGAGACCTTGAAGACTCCAGGAACCTTCAAGAGGCGCACGGA 1592
9046 CTGGCAGAGAAAAATAGACGAGGCGCTCGAAAGACTCCAGGAGCTTCAGGAAGCAACAGA 9105
1593 TGAGCTGGACCTCAAGCTGCGGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 1652
9106 TGAGCTGGATCTCAAACTACGTGAGGAGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 9165
1653 CGATCTCCTCATTGACTCTCTCCAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 1712
9166 TGACCTCCTCATTGACTCTCTCCAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 9225
1713 AATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCAC 1772
9226 AATTACACCTCTGAAAGAGAAATGTGACGTACGTCAATGACCTTGTCTGCCAAGCTTAC 9285
1773 TTTGGGCATTGAGCTCTACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATG 1832
9286 GTTGGGCATTGAGCTGTACCATATAACCTCAACACTCTGGAAGACCTGAACACCCAGATG 9345
1833 GAAGCTTCTGAGGTGGCGCTGAGGACCGGATCAGGAGCTGATGAAGCCCAAGCCACAGGGA 1892
9346 GAAGCTTCTGAGGTGGCGCTGAGGACCGGATCAGGAGCTGATGAAGCCCAAGCCACAGGGA 9405
1893 CTTTGGTCCAGCATCTCAGCACTTCTTCCAGCTCTGTCCAGGGTCCCTGGGAGAGAGC 1952
9406 CTTTGGACCAAGCCTCCAGCACTTCTTCCAGCTTCTGTCCAGGGTCCCTGGGAGAGAGC 9465
1953 CATCTCGCCAAACAAAGTGGCCCTACTATATCAACACGAGACTCAACA 2001
9466 CATCTCAACAAACAAAGTGGCCCTACTATATCAACACGAGACCCAAACA 9514

RESULT 14
AX306153 LOCUS AX306153 13815 bp DNA linear PAT 11-DEC-2001
DEFINITION Sequence 904 from Patent WO0188188.
ACCESSION AX306153
VERSION AX306153.1 GI:17645441
KEYWORDS Mus musculus (house mouse)
SOURCE

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
TITLE Method for examining ischemic conditions
JOURNAL Patent: WO 0188188-A 904 22-NOV-2001;
School Juridical Person Nihon University (JP)
FEATURES
Location/Qualifiers
1..13815
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
BASE COUNT 4577 a 2719 c 3025 g 3494 t
ORIGIN
Query Match 50.2%; Score 1004; DB 6; Length 13815;
Best Local Similarity 90.3%; Pred. No. 5.8e-241;
Matches 1073; Conservative 0; Mismatches 115; Indels 0; Gaps 0;
QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCCTCCCTGACCTGGAAAAAGTTTCTTGCTG 872
Db 8251 GGAAGAAACTCATAGATTACTGCAACAGTTCCTCCCTGACCTGGAGAAAGTTTCTTGCTG 8310
QY 873 GCTTACACAAGCTGAAACAACTGCCAATGTCTTACAGATGTCTACCGTAAGGAAAGCT 932
Db 8311 GATTACGGAAGCAGAAACAACTGCCAATGTCTTACAGGAGCTTCCGTAAGGAGAGCT 8370
QY 933 CCTAGAGACTCCAAAGGGAGTAAAGAGCTGATGAAACAAATGGCAAGACCTCCAAGGTGA 992
Db 8371 CCTAGAGACTCCAAAGGGAGTCAAGAGAGCTGATGAAACAAATGGCAAGATCTCCAAGGAGA 8430
QY 993 AATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGATAAAGCAGCCAAAAATCCTGAG 1052
Db 8431 AATTGAAGCTCACACAGATATCTATCACAACTCTTGATGATAAATGGCAAAAAATCCTGAG 8490
QY 1053 ATCCCTGGAAGGTTCCGATGATGCACTCCTGTTTACAAAGACGTTTGGATAACATGAACCT 1112
Db 8491 ATCCCTGGAAGGTTCCGATGATGAGCAACCCCTGTTTACAAAGACGTTTGGATAACATGAATTT 8550
QY 1113 CAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGTCCCATTTTGGAAAGCCAGTTC 1172
Db 8551 CAAGTGGAGTGAACCTTCAGAAAAAGTCTCTCAACATTAGTCCCATTTTGGAAAGCAAGTTC 8610
QY 1173 TGACCATGTGAAGCGTCTGCACCTTCTCTGACGAAACTTCTGTTGGCTACAGCTGAA 1232
Db 8611 TGACCATGTGAAGCGTCTGCATCTTCTCTTCAAGAACTTCTGTTGGCTACAGCTGAA 8670
QY 1233 AGATGATGAATTAAGCCGAGGACCTATTGGAGGCGACTTTCAGAGCAGTTTCAGAAAGCA 1292
Db 8671 AGATGATGAATTAAGCCGAGGACCTATTGGAGGCGACTTTCAGAGCAGTTTCAGAAAGCA 8730
QY 1293 GAACGATGTACATAGGCGCTTCAAGAGGGAATTTGAAAACTAAAGAACCTGTGAATCATGAG 1352
Db 8731 GAATGATATACATAGGCGCTTCAAGAGGGAATTTGAAAACTAAAGAACCTGTGAATCATGAG 8790
QY 1353 TACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCCTTTGGAAAGGACTAGAGAAACT 1412
Db 8791 TACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCCTTTGGAAAGGACTAGAGAAACT 8850
QY 1413 CTACAGGAGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT 1472
Db 8851 CTACAGGAGCCAGAGAGCTGCTCTCTGAGAGAGAGCTCAGAAATGTCACTCGGCTCCT 8910
QY 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGAACCTGCACTCCGCTGA 1532
Db 8911 ACGAAAGCAGGCTGAGGAGGTCAACGCTGAATGGGACAAATTTGAACCTGCGCTCAGCTGA 8970
QY 1533 CTGGCAGAGAAAAATAGATGAGACCTTGAAGAGACTCCAGGAACTTCAAGAGGCGCACGGA 1592
Db 8971 TTGGCAGAGAAAAATAGATGAAGCTCTTGAAGAGACTCCAGGAACTTCAAGAGAGCTGCCGA 9030
QY 1593 TGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGAATCTGGCAGCCCGTGGG 1652

Db 9031 TGAACCTGGACCTCAAGTTGCGCAAGCTGAGGTGATCAAGGATCCTGGCAGCCAGTGGG 9090
QY 1653 CGATCTCCTCATTTGACTCTCTCAAGATCACCTCGAAGAAAGTCAAGGCACCTTCGAGGAGA 1712
Db 9091 GGATCTCCTCATTTGACTCTCTGCAAGATCACCTTGAAAGAAAGTCAAGGCACCTTCGGGGAGA 9150
QY 1713 AATTGGGCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC 1772
Db 9151 AATTGCACTCTTAAGAGAAATGTCATCGTGTCAATGACCTTGACATCAGTGAACAC 9210
QY 1773 TTTGGGCAATTCAGCTCTCAACCGTATACCTCAGCACTTCGGAAGACCTGAACACAGATG 1832
Db 9211 ACTGGCAATTCAGCTCTCACCTTATAACCTCAGCACTTTGGAAGATCTGAATACCAGATG 9270
QY 1833 GAAGCTTCTGAGGTGGCCGCTGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGA 1892
Db 9271 GAGGCTTCTACAGGTGGCTGTGGAGGACCGTGTGACAGAGCTGCATGAAGCCACAGGGA 9330
QY 1893 CTTTGGTCCAGATCTCAGCACTTTCTTCCACGTCTGTCCAGGTCCCTGGGAGAGAGC 1952
Db 9331 CTTTGGTCTGCATCCAGCACTTCTTCCACTTCAGTTTCAAGGTCCCTGGGAGAGAGC 9390
QY 1953 CATCTCGCCAAACAAAGTGCCCTACTATATCAACACAGAGACTCAAAAC 2000
Db 9391 CATCTCACCAACAAAGTGCCCTACTATATCAACACAGAGACCCAAAC 9438
RESULT 15
AX538582 AX538582 13815 bp DNA linear PAT 23-NOV-2002
LOCUS
DEFINITION Sequence 2 from Patent WO0229056.
ACCESSION AX538582
VERSION AX538582.1 GI:25271088
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 2 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
1..13815
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
BASE COUNT 4577 a 2719 c 3025 g 3494 t
ORIGIN
Query Match 50.2%; Score 1004; DB 6; Length 13815;
Best Local Similarity 90.3%; Pred. No. 5.8e-241;
Matches 1073; Conservative 0; Mismatches 115; Indels 0; Gaps 0;
QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCCCTGACCTGGAAAGTTTCTTGCCTG 872
Db 8251 GGAAGAACTCATAGATTACTGAGCAGTTCCCTCTGGACCTGGAGAGTTTCTTCTCTG 8310
QY 873 GCTTACAGAACTGAAACAACTGCCAATGTCTACAGAGTGTACCCGTGAAGAAAGGCT 932
Db 8311 GATTACGGAAGCAGAAACAACTGCCAATGTCTACAGGACGCTTCCCGTAAGAGAGAGCT 8370
QY 933 CCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAACAAATGGCAAGACCTCAAGGTGA 992
Db 8371 CCTAGAAGACTCCAGGGAGTCAGAGAGCTGATGAACCAATGGCAAGATCTCCAAGGAGA 8430
QY 993 AATTGAAGCTCACACAGATGTTTATCAACACCTGGATGAACACAGCCAAAATCTGTAG 1052
Db 8431 AATTGAAGCTCACACAGATATCTATCACAATCTTGATGAAATGGCCAAAATCTGTAG 8490
QY 1053 ATCCCTGGAAGGTTCCGATGATGCAGTCCCTGTACAAAGACGTTTGGATAACATGAAC 1112

Db 8491 ATCCCTGGAAGGTTCCGATGAAGCACCCCTGTATCAAAAGACGTTTGGATAACATGAATTT 8550
QY 1113 CAAGTGGAGTGAACTTCGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAAGCAAGTTC 1172
Db 8551 CAAGTGGAGTGAACTTCGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAAGCAAGTTC 8610
QY 1173 TGACCAAGTGAAGCGCTCTGACCTTTCTCTGACGGAAGTCTCTGTTGGTGTACAGTGAA 1232
Db 8611 TGACCAAGTGAAGCGTTTGCATCTTTCTCTTCAAGGAATCTCTGTTTGGTGTACAGTGAA 8670
QY 1233 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACCTTTCCAGCAGTTTCAAGAGCA 1292
Db 8671 AGATGATGAATTAAGCGCTCAGGCACCCATCGGTGGTGAATTTCCAGCAGTTTCAAGAGCA 8730
QY 1293 GAACGATGTACATAGGGCCCTTCAAGAGGGAATTTGAAAACTAAAGAACCTGTATCATGAG 1352
Db 8731 GAATGATATACATAGGGCCCTTCAAGAGGGAATTTGAAAACTAAAGAACCTGTATCATGAG 8790
QY 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT 1412
Db 8791 TACTCTTGAGACTGTGAGAAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT 8850
QY 1413 CTACCAAGGAGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT 1472
Db 8851 CTACCAAGGAGCCAGAGAACTGCTCTCTGAGGAGAGAGCTCAGAAATGTCACTCGGCTCCT 8910
QY 1473 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGAACCTGCACCTCCGCTGA 1532
Db 8911 ACGAAAGCAGGCTGAAAGAGGTCAACGCTGAATGGGCAAAATTTGAACCTGCCTCAGCTGA 8970
QY 1533 CTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAATTTCAAGAGGCCACCGGA 1592
Db 8971 TTGGCAGAGAAAAATAGATGAAGCTCTTGAAGACTCCAGGAATTTCAAGAGGCCACCGGA 9030
QY 1593 TGAGCTGGACCTCAAGCTCGCCAAAGCTGAGGTGATCAAGGGATCTTGGCAGCCCGTGGG 1652
Db 9031 TGAACCTGGACCTCAAGTTGCGCCAAAGCTGAGGTGATCAAGGGATCTTGGCAGCCAGTGGG 9090
QY 1653 CGATCTCCTCATTTGACTCTCTGCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGGGGAGA 1712
Db 9091 GGATCTCCTCATTTGACTCTCTGCAAGATCACCTTTGAAAAAGTCAAGGCACCTTCGGGGAGA 9150
QY 1713 AATTGGGCTCTGAAAGAGAAAGTGAAGCCAGCTCAATGACCTTGTCTGCCAGCTTACCAC 1772
Db 9151 AATTGCACCTCTTAAAGAGAAATGTCAATCGTGTCAATGACCTTGCACATCAGCTGACCAC 9210
QY 1773 TTTGGGCATTCAGCTCTCACCGTATACCTCAGCACTCTGGAAGACTGAAACACCCAGATG 1832
Db 9211 ACTGGGCATTCAGCTCTCACCTTATAACCTCAGCACTTTGGAAGATCTGAATACCAGATG 9270
QY 1833 GAACTTCTGACAGGTGCGCTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGA 1892
Db 9271 GAGGCTTCTACAGGTGCTGTGGAGGACCGGTGTGACAGCAGCTGCATGAAGCCACAGGGA 9330
QY 1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTTCTGTCCAGGTCCCTGGGAGAGAGC 1952
Db 9331 CTTTGGTCCGATCCAGCACTTCTTTTCCACTTCAGTTTCAAGGTCCCTGGGAGAGAGC 9390
QY 1953 CATCTCGCCAAACAAAGTGCCCTACTATATCAACACAGAGACTCAAAAC 2000
Db 9391 CATCTCACCAACAAAGTGCCCTACTATATCAACACAGAGACCCAAAC 9438

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

MM nucleic - nucleic search, using sw model
Run on: February 1, 2004, 11:04:35 ; Search time 3939.6 Seconds
(without alignments)
12961.661 Million cell updates/sec

Title: US-09-845-416-2_COPY_900_3000
Perfect score: 2101
Sequence: 1 gagctatgcctacacacagg.....acctcagcactctggaagac 2101

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

| Database : | | | | EST:* | |
|------------|------------|---|--|-------|-------------|
| 1: | em_estba: | * | | 1: | em_estom: |
| 2: | em_esthum: | * | | 17: | em_gss_hum: |
| 3: | em_estin: | * | | 18: | em_gss_inv: |
| 4: | em_estmu: | * | | 19: | em_gss_pln: |
| 5: | em_estov: | * | | 20: | em_gss_vrt: |
| 6: | em_estpl: | * | | 21: | em_gss_fun: |
| 7: | em_estro: | * | | 22: | em_gss_mam: |
| 8: | em_htc: | * | | 23: | em_gss_mus: |
| 9: | gb_est1: | * | | 24: | em_gss_pro: |
| 10: | gb_est2: | * | | 25: | em_gss_rod: |
| 11: | gb_htc: | * | | 26: | em_gss_phg: |
| 12: | gb_est3: | * | | 27: | em_gss_vrl: |
| 13: | gb_est4: | * | | 28: | gb_gss1: |
| 14: | gb_est5: | * | | 29: | gb_gss2: |
| 15: | em_estfun: | * | | | |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| SUMMARIES | | | |
|------------|--------|-------------|----------------------------|
| Result No. | Score | Query Match | Description |
| 1 | 1103.4 | 52.5 | 3870 11 BC036103 Homo sapi |
| 2 | 866.6 | 41.2 | 3056 11 AK044536 Mus muscu |
| 3 | 597.2 | 28.4 | 728 14 CB228986 AGENCOURT |
| 4 | 592.4 | 28.2 | 834 12 BI729851 603349511 |

| | | | | | | |
|----|-------|------|------|----|----------|--------------------|
| 5 | 514.6 | 24.5 | 1490 | 11 | BC009242 | BC009242 Homo sapi |
| 6 | 509 | 24.2 | 595 | 14 | CB177816 | CB177816 is21c01.x |
| 7 | 491.6 | 23.4 | 750 | 12 | BI730168 | BI730168 603349711 |
| 8 | 486.6 | 23.2 | 579 | 9 | AL121550 | AL121550 DKFZp762L |
| 9 | 424.2 | 20.2 | 663 | 12 | BM488464 | BM488464 pgm2n.pk0 |
| 10 | 389 | 18.5 | 644 | 13 | BU313510 | BU313510 603540290 |
| 11 | 343 | 16.3 | 2334 | 11 | BC011062 | BC011062 Mus muscu |
| 12 | 324.6 | 15.4 | 402 | 14 | CB547284 | CB547284 AMGNNUC:S |
| 13 | 265.8 | 12.7 | 772 | 12 | BI250598 | BI250598 602993659 |
| 14 | 226 | 10.8 | 250 | 10 | BF963618 | BF963618 QV2-NN004 |
| 15 | 217.2 | 10.3 | 851 | 13 | BU201022 | BU201022 603952191 |
| 16 | 212.2 | 10.1 | 3051 | 11 | BC036095 | BC036095 Homo sapi |
| 17 | 203.6 | 9.7 | 1541 | 11 | AK034383 | AK034383 Mus muscu |
| 18 | 203.6 | 9.7 | 3753 | 11 | AK081426 | AK081426 Mus muscu |
| 19 | 192 | 9.1 | 778 | 14 | CB524596 | CB524596 UI-M-FY0- |
| 20 | 183 | 8.7 | 763 | 14 | CB518960 | CB518960 UI-M-GH0- |
| 21 | 182.4 | 8.7 | 645 | 9 | AW467977 | AW467977 he29g08.x |
| 22 | 177 | 8.4 | 646 | 12 | BI289102 | BI289102 UI-R-DK0- |
| 23 | 172.8 | 8.2 | 835 | 12 | BI553820 | BI553820 603190772 |
| 24 | 166.2 | 7.9 | 907 | 13 | BU122401 | BU122401 603003073 |
| 25 | 165.2 | 7.9 | 600 | 12 | BI988528 | BI988528 4012-24 M |
| 26 | 162 | 7.7 | 427 | 10 | BE817979 | BE817979 CM2-BN027 |
| 27 | 158.2 | 7.5 | 615 | 14 | CA377239 | CA377239 655768 NC |
| 28 | 153.2 | 7.3 | 885 | 11 | AK020881 | AK020881 Mus muscu |
| 29 | 150.6 | 7.2 | 515 | 28 | AZ780914 | AZ780914 2M0018010 |
| 30 | 144.8 | 6.9 | 434 | 13 | BQ375536 | BQ375536 PM1-TN012 |
| 31 | 144.6 | 6.9 | 675 | 14 | BY706879 | BY706879 BY706879 |
| 32 | 142.2 | 6.8 | 426 | 13 | BQ319056 | BQ319056 IL5-CT051 |
| 33 | 142.2 | 6.8 | 481 | 10 | BE557463 | BE557463 f103h06.Y |
| 34 | 141.6 | 6.7 | 794 | 28 | BZ097143 | BZ097143 CH230-236 |
| 35 | 141 | 6.7 | 761 | 28 | BZ129893 | BZ129893 CH230-452 |
| 36 | 140.6 | 6.7 | 649 | 13 | BQ209933 | BQ209933 UI-R-EP0- |
| 37 | 140 | 6.7 | 515 | 10 | BB651473 | BB651473 BB651473 |
| 38 | 137 | 6.5 | 466 | 14 | CD549993 | CD549993 B0305E01- |
| 39 | 134 | 6.4 | 828 | 28 | BZ126085 | BZ126085 CH230-452 |
| 40 | 133.2 | 6.3 | 521 | 28 | AZ226551 | AZ226551 RPCI-23-9 |
| 41 | 131.6 | 6.3 | 410 | 13 | BQ349936 | BQ349936 PM1-HT034 |
| 42 | 131.6 | 6.3 | 784 | 10 | BG212445 | BG212445 RST32032 |
| 43 | 129.2 | 6.1 | 545 | 13 | BX473852 | BX473852 DKFZp686J |
| 44 | 125.6 | 6.0 | 504 | 12 | BI289382 | BI289382 UI-R-DK0- |
| 45 | 124.4 | 5.9 | 681 | 13 | BU301653 | BU301653 603609005 |

ALIGNMENTS

| RESULT 1 | BC036103 | 3870 bp | mrna | linear | HTC 04-MAR-2003 |
|------------|---|--|------|--------|-----------------|
| BC036103 | BC036103 | Homo sapiens, clone IMAGE:5274415, mrna. | | | |
| LOCUS | BC036103 | | | | |
| DEFINITION | BC036103 | | | | |
| ACCESSION | BC036103.1 | GI:23271310 | | | |
| VERSION | HTC | | | | |
| KEYWORDS | Homo sapiens (human) | | | | |
| SOURCE | Homo sapiens | | | | |
| ORGANISM | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | |
| REFERENCE | 1 (bases 1 to 3870) | | | | |
| AUTHORS | Strausberg, R. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA | | | | |
| REMARK | NIH-MGC Project URL: http://mgc.nci.nih.gov | | | | |
| COMMENT | Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN) cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 | | | | |

Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 48 Row: f Column: 7
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein

This clone has the following problem: retained intron.

FEATURES

source
Location/Qualifiers
1..3870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5274415"
/tissue type="Brain, hippocampus"
/clone_lib="NIH MGC_95"
/lab_host="DH10B"
/note="Vector: pBluescript"
BASE COUNT 1321 a 677 c 763 g 1109 t
ORIGIN

Query Match 52.5%; Score 1103.4; DB 11; Length 3870;
Best Local Similarity 99.9%; Pred. No. 1.2e-209;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

| | | | |
|----|------|--|------|
| Qy | 1 | GAGCTATGCCCTACACACAGCGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAT | 60 |
| Db | 1043 | GAGCTATGCCCTACACACAGCGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAT | 1102 |
| Qy | 61 | TCCTTCACAGCATTTGGAGCTCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG | 120 |
| Db | 1103 | TCCTTCACAGCATTTGGAGCTCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG | 1162 |
| Qy | 121 | TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTC | 180 |
| Db | 1163 | TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTC | 1222 |
| Qy | 181 | TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA | 240 |
| Db | 1223 | TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA | 1282 |
| Qy | 241 | CCAGTTTCATCTCATGAGGGTACATGATGGATTGACAGCCCATCAGGCGCGGTTGG | 300 |
| Db | 1283 | CCAGTTTCATCTCATGAGGGTACATGATGGATTGACAGCCCATCAGGCGCGGTTGG | 1342 |
| Qy | 301 | TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA | 360 |
| Db | 1343 | TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA | 1402 |
| Qy | 361 | AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCGAAGATGGGAATGCCTCAGGGTAGC | 420 |
| Db | 1403 | AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCGAAGATGGGAATGCCTCAGGGTAGC | 1462 |
| Qy | 421 | TAGCATGGAATAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT | 480 |
| Db | 1463 | TAGCATGGAATAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT | 1522 |
| Qy | 481 | GAAAGATTGAATGACTGGCTTAACAAAACAGAGAAGAAACAAGGAAATGGAGGAAGA | 540 |
| Db | 1523 | GAAAGATTGAATGACTGGCTTAACAAAACAGAGAAGAAACAAGGAAATGGAGGAAGA | 1582 |
| Qy | 541 | GCCTCTTGGACCTGATCTTGAAGACCTAAACCGCCCAAGTACAACAATAGGTGCTTCA | 600 |
| Db | 1583 | GCCTCTTGGACCTGATCTTGAAGACCTAAACCGCCCAAGTACAACAATAGGTGCTTCA | 1642 |
| Qy | 601 | AGAAGATCTAGAAACAAGAACAGTCAGGGTCAATTTCTCACTCAGATGGTGGTAGT | 660 |
| Db | 1643 | AGAAGATCTAGAAACAAGAACAGTCAGGGTCAATTTCTCACTCAGATGGTGGTAGT | 1702 |

| | | | |
|----|------|--|------|
| Qy | 661 | TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGGAAGAACAACTTAAGGTATTGGG | 720 |
| Db | 1703 | TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGGAAGAACAACTTAAGGTATTGGG | 1762 |
| Qy | 721 | AGATCGATGGGCAAAACATCTGTAGATGGAGAGACCGCTGGTCTTTTACAAGACAT | 780 |
| Db | 1763 | AGATCGATGGGCAAAACATCTGTAGATGGAGAGACCGCTGGTCTTTTACAAGACAT | 1822 |
| Qy | 781 | CCTTCTCAAATGSCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTTACA | 840 |
| Db | 1823 | CCTTCTCAAATGSCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTTACA | 1882 |
| Qy | 841 | AAAAGAAGATGCAGTGAACAAGATTTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTT | 900 |
| Db | 1883 | AAAAGAAGATGCAGTGAACAAGATTTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTT | 1942 |
| Qy | 901 | ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT | 960 |
| Db | 1943 | ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT | 2002 |
| Qy | 961 | GGGCAAACTGTATTCACTCAAAACAGATCTTCTTTCAACACTGAAGATAAGTCAGTGAC | 1020 |
| Db | 2003 | GGGCAAACTGTATTCACTCAAAACAGATCTTCTTTCAACACTGAAGATAAGTCAGTGAC | 2062 |
| Qy | 1021 | CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTAGTCCAAA | 1080 |
| Db | 2063 | CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTAGTCCAAA | 2122 |
| Qy | 1081 | ACTTGAAAAGAGTACAGCACAGACT | 1105 |
| Db | 2123 | ACTTGAAAAGAGTACAGCACAGACT | 2147 |

RESULT 2
AK044536
LOCUS
DEFINITION
AK044536 3056 bp mRNA linear HTC 05-DEC-2002
Mus musculus adult retina cDNA, RIKEN full-length enriched library,
clone:A930019F21 product:dystrophin, muscular dystrophy, full
insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AK044536
AK044536.1 GI:26090404
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

REFERENCE

4 Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G., Tomita,M., Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Aono,H., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S. and Hayashizaki,Y.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409 (6821), 685-690 (2001)

MEDLINE 21085660

PUBMED 11217851

REFERENCE

AUTHORS

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3056)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoch,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gs.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Retina RNA was provided by Dr. Stefano Gustincich (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged. Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.

FEATURES

source

1. .3056
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:A930019F21"
/db_xref="taxon:10090"
/clone="A930019F21"
/tissue_type="retina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"

| | | |
|-----------------------|-------------------|---|
| misc_feature | 1. .3056 | /note="dystrophin, muscular dystrophy (MGD MG1:94909, GB NM_007868, evidence: BLASTN, 100%, match=999)" |
| BASE COUNT | 1072 a | 567 c 607 g 810 t |
| ORIGIN | | |
| Query Match | 41.2% | Score 866.6; DB 11; Length 3056; |
| Best Local Similarity | 86.5% | Pred. No. 1.9e-162; |
| Matches | 956; Conservative | 0; Mismatches 149; Indels 0; Gaps 0; |
| QY | 1 | GAGCTATGCCTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCAATT 60 |
| Db | 944 | GAGTTATGCCCTTCACACAGGCTGCTTATGTGGCACCTCTGATTCCACACAGAGCCCTTA 1003 |
| QY | 61 | TCCTTCACAGCATTTGGAAGCTCCTGAAGACAACTGTTGGCAGTTCAATTGATGGAGAG 120 |
| Db | 1004 | TCCTTCACAGCATTTGGAAGCTCCAGAGACAACTGACAGTTCACTTGACAGTTCAATTGATGGAGAC 1063 |
| QY | 121 | TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180 |
| Db | 1064 | GGAAGTAAATCTGGATAGTTACCAAACTGCTTTAGAAGAAGTACTTTTCATGGCTTCTTTC 1123 |
| QY | 181 | TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAATGATGTGGAAGTGGTAAAGA 240 |
| Db | 1124 | TGCCGAGGATACATTGGAGACACAAGGAGAGATTCAAATGATGTTGAAGAAGTAAAGA 1183 |
| QY | 241 | CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTCACAGCCCATCAGGCCCGGTTGG 300 |
| Db | 1184 | ACAGTTTTCATGCTCATGAGGGATTTCATGATGGATCTGACATCTCATCAAGGACTTGTGG 1243 |
| QY | 301 | TAATATTCTACAATTGGGAAGTAAGCTGATTGGGAACAGGAAAATTATCAGAAGATGAAGA 360 |
| Db | 1244 | TAATGTTTCTACAGTTAGGAAGTCAACTAGTTGGAAAAGGGAATTATCAGAAGATGAAGA 1303 |
| QY | 361 | AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420 |
| Db | 1304 | AGCTGAAGTGCAAGAACAAATGAATCTCCTAAATTCAAGATGGGAATGCTCAGGGTAGC 1363 |
| QY | 421 | TAGCATGGAAAAACAAGCAATTTCATAGAGTTTTAATGGATCTCCAGAAATCAGAAACT 480 |
| Db | 1364 | TAGCATGGAAAAACAAGCAAAATTACAAAAGTTCTAATGGATCTCCAGAAATCAGAAATT 1423 |
| QY | 481 | GAAGAGTTGAATGACTGGCTTAACAAAAACAGAAGAAAAGAACAGGAAAATGGAGGAAGA 540 |
| Db | 1424 | AAAAGAACTAGATGACTGGTTAAACAAAAACTGAAGAGAGAACTAAGAAAATGGAGGAAGA 1483 |
| QY | 541 | GCCTCTTGGACCTGATCTTGAAGACCTAAACCGCAAGTACAACAACATAAGGTGCTTCA 600 |
| Db | 1484 | GCCTCTTGGACCTGATCTTGAAGATCTAAATATGCCAAGTACAACAACATAAGGTGCTTCA 1543 |
| QY | 601 | AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTAGT 660 |
| Db | 1544 | AGAAGATCTAGAACAGGAGCAGGTCAAGGTCACTCGCTCACTCACATGGTAGTAGTGT 1603 |
| QY | 661 | TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTACTGGG 720 |
| Db | 1604 | TGATGAATCCAGCGGTGATCATGCAACAGCTGCTTTGGAAAGAACAACTTAAGGTACTGGG 1663 |
| QY | 721 | AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGGTGGTTCTTTTACAAGACAT 780 |
| Db | 1664 | AGATCGATGGGCAAAATATCTGCAGATGGACTGAAGACCGGTGGATTGTTTACAAGATAT 1723 |
| QY | 781 | CTTCTCAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTTCAGA 840 |
| Db | 1724 | CTTCTCAAATGGCAGCATTTTACTGAAGAACAGTGCCTTTTGTAGTACATGGCTTTTCAGA 1783 |
| QY | 841 | AAAAGAAGATGCAGTGAACAAGATTTCACACAATGGCTTTTAAAGATCAAAATGAAATGTT 900 |
| Db | 1784 | AAAAGAAGATGCAATGAAGAACATTTCAGACAAGTGGCTTTTAAAGATCAAAATGAAATGAT 1843 |
| QY | 901 | ATCAAGTCTTCAAAAAGTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 960 |
| Db | 1844 | GTCAAGTCTTTCACAAAATATCTACTTTTAAATAATAGATCTAGAAAAAGAAAAAGCAACCAT 1903 |

QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGATAAGTCAGTGAC 1020
|||
Db 1904 GGAATAAACTAAGTTCACTCAATCAAGATCTACTTTGGCCTGAAATAAATAGTCAGTGAC 1963
|||
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCGGTGTGGGATAATTTAGTCCAAA 1080
|||
Db 1964 TCAAAAGATGGAATCTGGATGGAATACTTTGCAACAAGTTGGGACAAATTAACCCAAA 2023
|||
QY 1081 ACTGAAAAGAGTACAGCACAGACT 1105
|||
Db 2024 ACTGAAAAGAGTTTCAGCACAAATT 2048
|||
RESULT 3
CB228986 728 bp mRNA linear EST 10-FEB-2003
LOCUS
DEFINITION AGENCOURT 11499247 NICHD Rh_Ov1 Macaca mulatta cDNA clone
IMAGE:6884820 5', mRNA sequence.
ACCESSION CB228986
VERSION
KEYWORDS
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
Cercopithecinae; Macaca.
REFERENCE 1 (bases 1 to 728)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Eliot Spindel
CDNA Library Preparation: CLONTECH
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM3135 row: e column: 11
High quality sequence stop: 583.
Location/Qualifiers
1..728
/organism="Macaca mulatta"
/mol_type="mRNA"
/db_xref="taxon:9544"
/clone="IMAGE:6884820"
/tissue_type="Ovary"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD Rh_Ov1"
/note="Organ: ovary; Vector: pDNR-LIB; Site 1: Sfi I;
Site 2: Sfi I; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.0-4.0 kb. Tissue pooled from
pre-pubertal, post pubertal sn menopausal monkeys.
Constructed by Clontech. Note: this is a NICHD Library."
BASE COUNT 256 a 127 c 162 g 182 t 1 others
ORIGIN
Query Match 28.4%; Score 597.2; DB 14; Length 728;
Best Local Similarity 97.0%; Pred. No. 8.6e-109;
Matches 619; Conservative 0; Mismatches 18; Indels 1; Gaps 1;
QY 295 GGTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATATCAGAAGA 354
|||
Db 1 GGTGGTAATATTCTACAATTGGGAAGTCAGCTGATTGGAACAGGGAATATCAGAAGA 60
|||
QY 355 TGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAG 414
|||
Db 61 TGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAG 120
|||
QY 415 GGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCA 474
|||

Db 121 GGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCA 180
|||
QY 475 GAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAAACAAGGAAATGGA 534
|||
Db 181 GAAACTGAAAGAGTTGAATGACTGGCTGACAAAAACAGAAGAAAGAAACAAGGAAATGGA 240
|||
QY 535 GGAAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACAATAAGGT 594
|||
Db 241 GAAAGAACCCCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACAATAAGGT 300
|||
QY 595 GCTTCAAGAAGATCTAGAACAAAGAACAGTCAGGGTCAATTCTCTCACTCACATGGTGGT 654
|||
Db 301 GCTTCAAGAAGATCTAGAACAAAGAACAGTCAGGGTCAATTCTCTCACTCACATGGTGGT 360
|||
QY 655 GGTAGTTGATGAATCTAGTGGAGATCACGGCACTGCTTTGGAGAAACAACCTTAAGGT 714
|||
Db 361 GGTAGTTGATGAATCTAGTGGAGATCACGGCACTGCTTTGGAGAGCAACTTAAGGT 420
|||
QY 715 ATTGGGAGATCGATGGCAAAACATCTGTAGTGGACAGAAGACCGCTGGGTTCTTTTACA 774
|||
Db 421 ATTGGGAGATCGATGGCAAAACATCTGCAGATGGACAGAAGACCGCTGGGTTCTTTTACA 480
|||
QY 775 AGACATCCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCT 834
|||
Db 481 AGACATCCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCT 540
|||
QY 835 TTCAGAAAAAGAGATGAGTGAAGTGAACAAAGATTACACAACT-GGCTTTAAAGATCAAAATG 893
|||
Db 541 TTCAGAAAAAGAGATGAGTGAAGTGAACAAAGATTACACAACTGGGCTTTAAAGGATCAAAATG 600
|||
QY 894 AAATGTTATCAAGTCTTCAAAAAACTGGCCGTTTAAAAA 931
|||
Db 601 AAATGTTATCAAGTCTTCAAAAAACTGGCCCGTATGTAA 638
|||
RESULT 4
BI729851
LOCUS
DEFINITION BI729851 NIH_MGC_94 Mus musculus cDNA clone linear EST 20-SEP-2001
603349511F1
mRNA sequence.
ACCESSION BI729851
VERSION BI729851.1 GI:15706864
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 834)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11908 row: e column: 03
High quality sequence stop: 796.
Location/Qualifiers
1..834
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:5357162"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 94"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Source

Qy 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480
|||||
Db 1389 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 1448
|||||
Qy 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAAGAAAGAA 521
|||||
Db 1449 GAAAGAGTTGAATGACTGGCTAACAAAAACAAGAAAGAA 1489
|||||
RESULT 6
CB177816/c
LOCUS
DEFINITION is21c01.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6553129 3'
similar to SW:DMD_HUMAN P11532 DYSTROPHIN. [1] ; mRNA sequence.
CB177816
CB177816.1 GI:28186206
EST.
Homo sapiens (human)
Homo sapiens
ORGANISM
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Sceaare,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.,
Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished
COMMENT Other-ESTs: is21c01.y1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 448.
Location/Qualifiers
1. .595
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6553129"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_lib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
Size-selected on agarose gel. Average insert size ~1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."
BASE COUNT 146 a 136 c 116 g 197 t
ORIGIN
Query Match 24.2%; Score 509; DB 14; Length 595;
Best Local Similarity 100.0%; Pred. No. 3.3e-91;
Matches 509; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GAGCTATGCTACACAGGCTGCTTATGTCACCACCTTGACCCCTACACGGAGCCCAT 60
|||||

Db 509 GAGCTATGCTACACACAGGCTGCTTATGTCACCACCTCTGACCTACACGGAGCCCAT 450
Qy 61 TCCTTCACAGCAATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG 120
|||||
Db 449 TCCTTCACAGCAATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG 390
|||||
Qy 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
|||||
Db 389 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 330
|||||
Qy 181 TGCTGAGGACACATTCGAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
|||||
Db 329 TGCTGAGGACACATTCGAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 270
|||||
Qy 241 CCAGTTTCATCTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGCCCGGGTTGG 300
|||||
Db 269 CCAGTTTCATCTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGCCCGGGTTGG 210
|||||
Qy 301 TAATATTCTACAATTGGGAAGTAAGCTGATGTTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
|||||
Db 209 TAATATTCTACAATTGGGAAGTAAGCTGATGTTGGAACAGGAAAAATTATCAGAAGATGAAGA 150
|||||
Qy 361 AACTGAAGTACAAGACAGCAGATGAATCTCTTAAATTCGAAGATGGGAATGCCTCAGGGTAGC 420
|||||
Db 149 AACTGAAGTACAAGACAGCAGATGAATCTCTTAAATTCGAAGATGGGAATGCCTCAGGGTAGC 90
|||||
Qy 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480
|||||
Db 89 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 30
|||||
Qy 481 GAAAGAGTTGAATGACTGGCTAACAAAA 509
|||||
Db 29 GAAAGAGTTGAATGACTGGCTAACAAAA 1
|||||
RESULT 7
BI730168
LOCUS
DEFINITION BI730168 750 bp mRNA linear EST 20-SEP-2001
603349711F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5357187 5',
mRNA sequence.
ACCESSION BI730168
VERSION BI730168.1 GI:15707181
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 750)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM11908 Row: f column: 04
High quality sequence stop: 747.
Location/Qualifiers
1. .750
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:5357187"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_94"
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed."
FEATURES
source

Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

| | | | | | |
|-----------------------|-----------------|--|-----------|-------------|----------|
| BASE COUNT | 240 a | 161 c | 178 g | 170 t | 1 others |
| ORIGIN | | | | | |
| Query Match | 23.4%; | Score 491.6; | DB 12; | Length 750; | |
| Best Local Similarity | 89.3%; | Pred. No. 9.9e-88; | | | |
| Matches 598; | Conservative 0; | Mismatches 64; | Indels 8; | Gaps 6; | |
| QY | 1098 | CACAGACTCATAGATTACTGCAACAGTTCCCTGGACCTGGAAAAAGTTTCTTGCTGG- | 1156 | | |
| Db | 80 | CAGAACTCATAGATTACTGCAGCAGTCCCTCTGGACCTGGAGAAAGTTTCTTCTCTGGA | 139 | | |
| QY | 1157 | CTTACAGAACTGAAAC--AACTGCCAATGTCTCTACAGGATGCTACCCG-TAAGGAAAGG | 1213 | | |
| Db | 140 | TTTACGGAAGCAGAAACAGAGCTGCCAATGTCTCTACAGGACGCTTCCCGCTAAGGAGAAG | 199 | | |
| QY | 1214 | CTCCTAGAAAGACTCCAAGGG--AGTAAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAG | 1271 | | |
| Db | 200 | CTCCTAGCAAGACTCCAGGGGCAGTCAGAGAGCTGATGAAACCATGGCAAGATCTCCAAG | 259 | | |
| QY | 1272 | GTGAAA-TTGAAGCTCACACAGATGTTTATCAACACCTGGATGAAACAGCCRAAAAATC | 1330 | | |
| Db | 260 | GAGAAAGTTGAAACTCACACAGATATCTATCAATCTTGATGAAATGGCCRAAAAATC | 319 | | |
| QY | 1331 | CTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTACAAAGACGTTTGGATAACATG | 1390 | | |
| Db | 320 | CTGAGATCCCTGGAAGGTTCCGATGAAGCACCCCTGTTACAAAGACGTTTGGATAACATG | 379 | | |
| QY | 1391 | AACCTCAAGTGGAGTGAACCTTCGGA AAAAGTCTCTCAACATTAGTCCCATTTGGAAAGCC | 1450 | | |
| Db | 380 | AATTTCAAGTGGAGTGAACCTTCAGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAAGCA | 439 | | |
| QY | 1451 | AGTTCTGACCAGTGAAGCGTCTGCACCTTTCTCTGCGAGAACTTCTGGTGTGGCTACAG | 1510 | | |
| Db | 440 | AGTTCTGACCAGTGAAGCGTTTGGCATCTTTCTCTTCAGGAACTTCTTGTGGCTACAG | 499 | | |
| QY | 1511 | CTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGCGGACTTTCAGCAGTTCAG | 1570 | | |
| Db | 500 | CTGAAAGATGATGAACCTGAGCCGTCAGGCACCCATCGGTGGTGATTTCCAGCAGTTCAG | 559 | | |
| QY | 1571 | AAGCA-GAAGCATGTACATAGGCGCTTCAAGAGGGGAATTAAGAACTAAAGAACCTGTAAAT | 1629 | | |
| Db | 560 | AAGCAGCAATGATATACATAGGCGCTTCAAGAGGGGAATTAAGAACTAAAGAACCTGTAAAT | 619 | | |
| QY | 1630 | CATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCGCTTTGGAGGACTAGA | 1689 | | |
| Db | 620 | CATGAGTACTCTTGAGACTGTGAGAATATTCTGACAGAGCAGCGCTTTGGAGGACTAGA | 679 | | |
| QY | 1690 | GAAACTCTACCAGGAGCCAGAGAGTGCCTCCTGAGGAGAGAGCCAGAAATGTCACTCG | 1749 | | |
| Db | 680 | GAAACTCTACCAGGAGCCAGAGAGTGCCTCCTGAGGAAAGAGCTCAGAAATGTCACTCG | 739 | | |
| QY | 1750 | GCTTCTACGA | 1759 | | |
| Db | 740 | GCTCCTACGA | 749 | | |

RESULT 8
AL121550
LOCUS
DEFINITION
DKFZp762L078 r1 762 (synonym: hmcl2) Homo sapiens cDNA clone
DKFZp762L078 5', mRNA sequence.

ACCESSION
AL121550
VERSION
AL121550.1 GI:5927551
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 579)
AUTHORS
Ottenwaelder,B., Obermaier,B., Mewes,H.W., Gassenhuber,J. and

Wiemann,S.
EST (Ottenwaelder, et al.)
Unpublished
Contact: Ottenwaelder B
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by MediGenomix (Martinsried/Germany) within the cDNA
sequencing consortium of the German Genome Project. No sl sequence
available.
This clone (DKFZp762L078) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
1..579
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp762L078"
/tissue_type="melanoma (MeWo cell line)"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="762 (synonym: hmcl2)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

BASE COUNT 180 a 128 c 144 g 127 t
ORIGIN
Query Match 23.2%; Score 486.6; DB 9; Length 579;
Best Local Similarity 97.2%; Pred. No. 9.5e-87;
Matches 495; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

| | | | | | |
|----|------|---|------|--|--|
| QY | 1260 | AAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAAACTGGATGAAACA | 1319 | | |
| Db | 71 | AGGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAAACTGGATGAAACA | 130 | | |
| QY | 1320 | GCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGAGTCTCTTACAAAGACGTT | 1379 | | |
| Db | 131 | GCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGAGTCTCTTACAAAGACGTT | 190 | | |
| QY | 1380 | TGGATAACATGAACCTCAAGTGAGTGAACCTTCGGA AAAAGTCTCTCAACATTAGTCC | 1439 | | |
| Db | 191 | TGGATAACATGAACCTCAAGTGAGTGAACCTTCGGA AAAAGTCTCTCAACATTAGTCC | 250 | | |
| QY | 1440 | ATTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTCTCTGCAGGAACCTCTGG | 1499 | | |
| Db | 251 | ATTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTCTCTGCAGGAACCTCTGG | 310 | | |
| QY | 1500 | TGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTC | 1559 | | |
| Db | 311 | TGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTC | 370 | | |
| QY | 1560 | CAGCAGTTCAGAAAGCAGACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACATAAG | 1619 | | |
| Db | 371 | CAGCAGTTCAGAAAGCAGACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACATAAG | 430 | | |
| QY | 1620 | AACCTGTAATCATGAGTACTTTGAGACTGTACGAAATTTCTGACAGAGCAGCCTTTGG | 1679 | | |
| Db | 431 | AACCTGTAATCATGAGTACTTTGAGACTGTACGAAATTTCTGACAGAGCAGCCTTTGG | 490 | | |
| QY | 1680 | AAGGACTAGAGAAAACCTCTACAGAGCCCAGAGAGTGCCTCCTGAGGAGAGAGCCAGAG | 1739 | | |
| Db | 491 | AAGGGCTAGAGAAAACCTCTACAGAGCCCAGAGAGTGCCTCCTGAGGAGAGAGCCAGAG | 550 | | |
| QY | 1740 | ATGTCACTCGGCTTCTACGAAAGCAGGCT | 1768 | | |
| Db | 551 | AATGTCACTCGCTTCTACGAAAGCAGGCT | 579 | | |

RESULT 9
BM488464
LOCUS
BM488464
663 bp mRNA linear EST 07-FEB-2002

DEFINITION pgm2n.pk007.117 Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library (pgm2n) Gallus gallus cDNA clone pgm2n.pk007.117 5' similar to sp|P11533|DMD CHICK Dystrophin p1r|S02041 dystrophin, muscle - chicken emb|CAA31746.1| (X13369) dystrophin (AA 1 - 3660) [Gallus gallus], mRNA sequence.

ACCESSION BM488464

VERSION BM488464.1 GI:18609395

KEYWORDS EST.

SOURCE Gallus gallus (chicken)

ORGANISM Gallus Gallus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

AUTHORS 1 (bases 1 to 663)

TITLE Cogburn,L.A. and Monsonogo-Ornan,E.

JOURNAL ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library, USDA/IFAFS Animal Genome Project

COMMENT Unpublished

Contact: Larry A. Cogburn

University of Delaware

Townsend Hall, Newark, DE 19717, USA

Tel: 302-831-1335

Fax: 302-831-2822

Email: cogburn@udel.edu, www.chickest.udel.edu.

FEATURES

source

1. .663

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="Commercial broiler and Ottawa Res. Centre Strains 90 & 21"

/db_xref="taxon:9031"

/clone="pgm2n.pk007.117"

/sex="Male and Female"

/tissue type="Breast muscle, leg muscle and epiphyseal growth plate"

/dev_stage="Breast,leg:Embryo(d19);post-hatch(1d,1,3,5,7,9,11 weeks);growth plate(1d,7d,14d post-hatch)"

/lab_host="E. coli EMDH10B"

/clone_lib="Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library (pgm2n)"

/note="Vector: pCMVSPORT6; Library made from equivalent pools of total RNA isolated from each tissue (embryonic muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth plate 33.3% of the final RNA pool). Single pass sequencing from 5'-end"

BASE COUNT 209 a 132 c 172 g 144 t 6 others

ORIGIN

Query Match 20.2%; Score 424.2; DB 12; Length 663;

Best Local Similarity 78.1%; Pred. No. 2.7e-74;

Matches 518; Conservative 0; Mismatches 144; Indels 1; Gaps 1;

QY 1243 GCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTATCA 1302

Db 1 GCTCATGAAGCAGTGGCAGGATCTACAGGCAGAAATTGATGCACATACTGACATCT-TNN 59

QY 1303 CAACCTGGATGAAACACAGCCAAATAATCCTGAGATCCCTGGAAGTTCCGATGATGCAGT 1362

Db 60 CAACCTGGATGAAACACGGGCAGAAATCCTGAGATCCCTGGAAGCTCAGAGGATGCTGT 119

QY 1363 CCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTCGGAAAAGTC 1422

Db 120 CCTGTTGCAGACGCTCTGGATAACATGAACCTTCAGATGGAGTGAGCTTAGGAAGAATC 179

QY 1423 TCTCAACATTAGGTCCTCCATTTGGAGCCAGTTCTTGACCAGTGAAGCGTCTGCACCTTTC 1482

Db 180 TCTAAACATTAGATCTCATTTTGGAGCCAGCACAGACCAGTGAAGCGTTTACATCTCTC 239

QY 1483 TCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACC 1542

Db 240 TCTTCAGGAACCTTTGGCATGGCTGCAATTGAAGGAGGATGAATTAACACAGCAGCACC 299

QY 1543 TATTGGAGGCGGACTTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAG 1602

Db 300 CATTTGGTGAGATATTTCCCACTGTGCAGAAGCAGAATGATGTTTCATAGGACTTTCAAGAG 359

QY 1603 GGAATTGAAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCT 1662

Db 360 GGAGCTGAAAAACAAAAGAACCTGTTATCATGAATGCACTTGAGACTGTGCGACTCTTCCT 419

QY 1663 GACAGAGCAGCCCTTTTGGAAAGGACTAGAGAAAACCTCTACCAGGAGCCAGAGAGCTGCCTCC 1722

Db 420 GGCAGATCAACCAGTAGAGGACTGGAAAAGGTCTATCCAGAACCAAGAGAGACTATCACC 479

QY 1723 TGAGGAGAGAGCCCGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATAC 1782

Db 480 TGAGGAGAGGGCCCGAATGTCACTAAAGTTCTCCGAAGGCAAGCAGATGATGTCAGNAC 539

QY 1783 TGAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCT 1842

Db 540 TGAGTGGGATAAGCTAATCTACGTTCTGCTGATTGGCAAAAGAGATAGATGATGCTCT 599

QY 1843 TGAAGACTCCAGGAACCTTCAAGAGGCCACCGATGATGAGCTCAAGCTGCGCCCAAGC 1902

Db 600 TGAAGACTGCAGGGTCTTCAGGAGGCAATGGATGAATNNNCTGAAACTGAACTGCGCCAGGC 659

QY 1903 TGA 1905

Db 660 TGA 662

RESULT 10

BU313510

LOCUS 603540290F1 CSEQCHN61 Gallus gallus cDNA clone ChEST508c24 5', mRNA

DEFINITION sequence.

ACCESSION BU313510

VERSION BU313510.1 GI:25821511

KEYWORDS EST.

SOURCE Gallus gallus (chicken)

ORGANISM Gallus Gallus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

AUTHORS 1 (bases 1 to 644)

Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E., Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.

A Comprehensive Collection of Chicken CDNAS

Curr. Biol. 12 (22), 1965-1969 (2002)

22335534

12445392

COMMENT Contact: Simon Hubbard

Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology (UMIST)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

location/Qualifiers

1. .644

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="Compton Line 151"

/db_xref="taxon:9031"

/clone="ChEST508c24"

/sex="Female"

/dev_stage="adult"

/lab_host="DH10B"

/clone_lib="CSEQCHN61"

/note="Organ: heart; Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction.

FEATURES

source

Following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used.

| | | | | | |
|-----------------------|-----------------|--|-----------|-------------|----------|
| BASE COUNT | 192 a | 141 c | 168 g | 142 t | 1 others |
| ORIGIN | | | | | |
| Query Match | 18.5%; | Score 389; | DB 13; | Length 644; | |
| Best Local Similarity | 75.5%; | Pred. No. 2.8e-67; | | | |
| Matches 482; | Conservative 0; | Mismatches 156; | Indels 0; | Gaps 0; | |
| | | | | | |
| y | 1449 | CCAGTTCTGACCAGTGAAGCGTCTGCACCTTCTCTGCAGGAACCTCTGGTGTGGCTAC | 1508 | | |
| b | 7 | CGAGCACAGACCAGTGAAGCGTTTACATCTCTCTCTCAGGAACCTTTGGCATGGCTGC | 66 | | |
| | | | | | |
| y | 1509 | AGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTC | 1568 | | |
| b | 67 | AATTGAAGGAGGATGAATTAACACAGCAAGCACCCATTGGTGGAGATATTCACACTGTGC | 126 | | |
| | | | | | |
| y | 1569 | AGAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACCTAAAGAACCTGTAA | 1628 | | |
| b | 127 | AGAAGCAGAAATGATGTTTCATAGGACTTTCAAGAGGGAGCTGAAAACAAAGAACCTGTTA | 186 | | |
| | | | | | |
| y | 1629 | TCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAG | 1688 | | |
| b | 187 | TCATGAATGCACCTTGACACTGTGCGACTCTTCCTGGCAGATCAACCACTAGAGGACTGG | 246 | | |
| | | | | | |
| y | 1689 | AGAAACTCTACAGGAGCCACAGAGAGCTGCCTCTGAGGAGAGAGCCCAAGATGTCACTC | 1748 | | |
| b | 247 | AAAAGGTCTATCCAGAACCAAGAGACCTATCACTGAGGAGAGGGCCCAAGATGTCACTA | 306 | | |
| | | | | | |
| y | 1749 | GGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACGTAGTGGGAAAATGAACCTGCAC | 1808 | | |
| b | 307 | AAGTTCTCCGAGGCAAGCAGATGATGTGAGAACTGAGTGGGATAAGCTAAATCTACGTT | 366 | | |
| | | | | | |
| y | 1809 | CCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGG | 1869 | | |
| b | 367 | CTGCTGATTGGCAAAAGAGATAGATGATGCTCTTGAAGACTGCAGGCTCTTCAGGAGG | 426 | | |
| | | | | | |
| y | 1869 | CCACGGATGAGCTGGACCTCAAGCTGCGCCAGCTGAGGTGATCAAGGATCCTGGCAGC | 1928 | | |
| b | 427 | CAATGGATGAACCTAGACCTGAAACTGCGCCAGGCTGAAGCATTCAGGGATCCTGGCAGC | 486 | | |
| | | | | | |
| y | 1929 | CCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTC | 1988 | | |
| b | 487 | CAGTGGGGATCTGCTGATAGACTCTCTGCAGGATCACTTAGAAAAGTCAAGGTTTATC | 546 | | |
| | | | | | |
| y | 1989 | GAGGAGAAATTGGCCCTCTGAAAGAGAACGCTGAGCCACGTCATATGACCTTGCTGCCCAGC | 2048 | | |
| b | 547 | GAGCAGAAATGGTGCCCTTANAGAGAAGGTGCTCAAGTCAATGAGCTGCTCACC | 606 | | |
| | | | | | |
| y | 2049 | TTACCACTTTGGGCATTTCAGTCTCACC | 2086 | | |
| b | 607 | TCGCTCCCTGATATTTCAGTTCTCCCATACACTCTC | 644 | | |

RESULT 11
LOCUS BC011062 2334 bp mRNA linear HTC 04-MAR-2003
DEFINITION Mus musculus, similar to utrophin, clone IMAGE:3979320, mRNA.
ACCESSION BC011062
VERSION BC011062.1 GI:15029695
KEYWORDS HTC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2334)

AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (25-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 23 Row: j Column: 4
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein
This clone has the following problem: retained intron.

FEATURES
Location/Qualifiers
1..2334
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3979320"
/tissue_type="Mammary tumor. Metallothionien-TGF alpha model. 10 month old virgin mouse. Taken by biopsy."
/clone_lib="NCI CGAP_Mam1"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6"

BASE COUNT 698 a 512 c 638 g 486 t
ORIGIN

Query Match 16.3%; Score 343; DB 11; Length 2334;
Best Local Similarity 59.0%; Pred. No. 5e-58;
Matches 589; Conservative 0; Mismatches 410; Indels 0; Gaps 0;

| | | | |
|----|------|---|------|
| QY | 104 | AGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAAGTA | 163 |
| Db | 1042 | AGCACCGTCACTGAAGTGGACATGGATTTGGACAGCTACCAGATAGCGCTAGAGGAAGTG | 1101 |
| | | | |
| QY | 164 | TTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACCAAGGAGAGATTCTAATGAT | 223 |
| Db | 1102 | CTGACGTGGCTGCTGTCCGCGGAGGACACGTTCCAGGAGCAAGATGACATTCTGATGAT | 1161 |
| | | | |
| QY | 224 | GTGGAAGTGGTAAAGACCCAGTTTCATACCTCATGAGGGGTACATGATGGATTTCACAGCC | 283 |
| Db | 1162 | GTGGAAGAAGTCAAGAGCAGTTTGCTACCCATGAAACTTTTATGATGGAGCTGACAGCA | 1221 |
| | | | |
| QY | 284 | CATCAGGGCCGGTTGGTAAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAA | 343 |
| Db | 1222 | CACCAGAGCAGCGTGGGAGCGTCTCTGCAGGCTGCAACCAAGCTGATGACACAAGGGA | 1281 |
| | | | |
| QY | 344 | TTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGG | 403 |
| Db | 1282 | CTGTACAGGAGGAGGAGTTTGAGATCCAGGAACAGATGACCTTGTCTGAATGCAAGGTGG | 1341 |
| | | | |
| QY | 404 | GAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGAT | 463 |
| Db | 1342 | GAGGCGCTCCGGTGGAGAGCATGGAGGCGAGTCCCGGCTGCACGCGCTCTGATGGAG | 1401 |
| | | | |
| QY | 464 | CTCCAGAATCAGAAACTGAAAGAGTTGAATGATCGCTAAACAAAAACAGAAAGAACA | 523 |

Db 1402 CTGCAGAAGAAACAGCTGCAGCAGCTCTCAAGCTGGCTGGCCCTCACAGAAGAGCGCAT 1461

QY 524 AGGAAAATGGAGAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAA 583

Db 1462 CAGAAGATGGAGAGCCTCCCGTGGTGATGACCTGCCCTCCCTGCAGAAAGCTGCTTCAA 1521

QY 584 CAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAGGTCATTTCTCTCACT 643

Db 1522 GAACATAAAAGTTTGCAAAATGACCTTGAAGTGAACAGGTGAAGTAAATTCCTTAAC 1581

QY 644 CACATGGTGGTGTAGTGTATGAATCTAGTGGAGATCACGCAACTGCTTGTGGAAGAA 703

Db 1582 CACATGGTGTGATTGTGGATGAAACACAGTGGGGAGAGTGCCACAGCTCTTCTGGAAGAT 1641

QY 704 CAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGCTGG 763

Db 1642 CAGTTACAGAAACTGGGTGAGCGCTGGACAGCTGTATGCCGTGGACTGAAGAACGTTGG 1701

QY 764 GTTCTTTTACAAGACATCCTTCTCAATATGGCAACGTCTTACTGAAGAACAGTGCCTTTT 823

Db 1702 AACAGGTTGCAAGAATCAGTATTCTGTGGCAGGAATTATTGGAAGACAGTGTCTGTTG 1761

QY 824 AGTGCATGGCTTTCAGAAAAGAAAGATGCAGTGAACAAGATTCACACAAGCTGGCTTTAA 883

Db 1762 GAGGCTTGGCTCACCGAAAAGGAAGAGGCTTTTGAATAAGTTCAAACAGCAACTTTAA 1821

QY 984 GATCAAAATGAAATGTTATCAAGTCTTCAAAAAGTGGCCGCTTTTAAACGGGATCTAGAA 943

Db 1822 GACCAGAAGGAACATAAGTGTCAAGTGTCCGGCGTCTGGCTATATTGAAGGAAGACATGGA 1881

QY 944 AAGAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTG 1003

Db 1882 ATGAAGAGGCAGACTCTGGATCAACTGAGTGAGATGGCCAGGATGTGGGCCAATTACTC 1941

QY 1004 AAGATAAGTCAGTGACCCAGAACGGAAGCATGGCTGGATAACTTTGGCCGGTGTGG 1063

Db 1942 AGTAATCCCAAGGCATCTAAGAAGATGAACAGTGAATCTGAGGAGCTAACACAGAGATGG 2001

QY 1064 GATAATTTAGTCCAAAAACTTCAAAAAGAGTACAGCACAG 1102

Db 2002 GATTCTCTGGTTCAGAGACTCGAAGACTCTTCTTAACAG 2040

RESULT 12
CB547284/c
LOCUS
DEFINITION
CB547284 402 bp mRNA linear EST 01-APR-2003
AMGNNUC:SRPB2-00242-E3-A srpb2 (10220) Rattus norvegicus cDNA clone
srpb2-00242-e3 5', mRNA sequence.

ACCESSION
CB547284
VERSION
CB547284.1 GI:29431225
KEYWORDS
EST.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 402)
AUTHORS
Amgen EST Program.
TITLE
Amgen Rat EST Program
JOURNAL
Unpublished
COMMENT
Contact: Dan Fitzpatrick
Amgen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00242 row: e column: 3.
Location/Qualifiers
1. .402
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="srpb2-00242-e3"
/tissue_type="prostate tissue"
/clone_lib="srpb2 (10220)"

/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; rat
prostate normalized double selected poly(A+) mRNA size
fraction > 1 kb"
BASE COUNT 83 a 111 c 88 g 119 t 1 others
ORIGIN

Query Match 15.4%; Score 324.6; DB 14; Length 402;
Best Local Similarity 90.8%; Pred. No. 1.9e-54;
Matches 345; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1519 TGATGAATTAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAAAGCAGAA 1578

Db 402 TGATGAGTTGAGCCGGCAGGCACCTATTGGTGGCGATTTTCCAGCAGTTTCAGAAAGCAGAA 343

QY 1579 CGATGTACATAGGCGCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAAATCATGAGTAC 1638

Db 342 TGATGTACACAGGCGCTTCAAGAGGGGAATTGANAACCTAAAGAACCTGTAAATCATGAGTAC 283

QY 1639 TCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAGGAGCTAGAGAACTCTA 1698

Db 282 TCTTGAGACTGTGAGAATATTTCTGACAGAGCAGCCCTTTGGAGGAGCTAGAGAACTCTA 223

QY 1699 CCAGGAGCCCGCAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAAATGTCACCTCGGCTTCTACG 1758

Db 222 CCAGGAGCCCGCAGAGAGCTGCCTCCTGAAAGAAAGAGCTCAGAATGTCACTCGGCTCCTACG 163

QY 1759 AAAGCAGGCTGAGGAGGTCAATATACTGAGTGGGAAAAAATTGAACCTCGACTCCGCTGACTG 1818

Db 162 AAAGCAGGCTGAAGAGGTCAACACTGAATGGACAAAATTGAACCTCGACTCAGCTGATTG 103

QY 1819 GCAGAGAAAAATAGATGAGAGCCCTTGAAGAGACTCCAGGAACCTTCAAGAGGCCACGGATGA 1878

Db 102 GCAGCGAAAAATAGATGAGGCTCTTTGAAGAGACTGCAGGAACCTTTCAGGAAGCTGCCGATGA 43

QY 1879 GCTGGACCTCAAGCTGCGCC 1898

Db 42 ACAGCACCTCATGTTGCGGC 23

RESULT 13
BI250598
LOCUS
DEFINITION
BI250598 772 bp mRNA linear EST 17-JUL-2001
602993659F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5149431 5',
mRNA sequence.

ACCESSION
BI250598
VERSION
BI250598.1 GI:14799101
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 772)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11369 row: e column: 16
High quality sequence start: 3
High quality sequence stop: 771.
Location/Qualifiers
1. .772
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"


```

BU201022
LOCUS
DEFINITION
    BU201022
    603952191P1 CSEQCHN03 Gallus gallus CDNA clone ChEST914e23 5', mRNA
    sequence.
    BU201022
    BU201022.1 GI:25364517
    EST.
    KEYWORDS
    SOURCE
    ORGANISM
        Gallus gallus (chicken)
        Gallus gallus
        Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
        Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
        Phasianinae; Gallus.
        1 (bases 1 to 851)
REFERENCE
    Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
    Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
    A Comprehensive Collection of Chicken cDNAs
    Curr. Biol. 12 (22), 1965-1969 (2002)
    22335534
    12445392
COMMENT
    Contact: Simon Hubbard
    Department of Biomolecular Sciences
    University of Manchester Institute of Science and Technology (UMIST
    )
    PO Box 88, Manchester, M60 1QD, UK
    Tel: 01612008930
    Fax: 01612360409
    Email: Simon.Hubbard@umist.ac.uk.
        Location/Qualifiers
        1..851
FEATURES
    source

```

| BASE COUNT | 269 a | 169 c | 238 g | 175 t | |
|---------------------------|--------|--|-----------|-------------|--|
| ORIGIN | | | | | |
| Query Match | 10.3%; | Score 217.2; | DB 13; | Length 851; | |
| Best Local Similarity | 61.8%; | Pred. No. 5.4e-33; | | | |
| Matches 362; Conservative | 0; | Mismatches 223; | Indels 1; | Gaps 1 | |
| QY | 347 | TCAGAAGATGAAGAAACTGAAGTACAAAGCAGATGAATCTCTAAATTCAAAGATGGAA | 406 | | |
| DB | 2 | TCATCTGAAGAGGAGTTTGAGATTTCAGGAACAAATGCTGCTGCTGAACCTCCGCTGGAG | 61 | | |
| QY | 407 | TGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTC | 466 | | |
| DB | 62 | GACCTCAGGTTGAGAGCATGGACAGGCACTCCCGCTGCATGACATGCTGATGGAGCTA | 121 | | |
| QY | 467 | CAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAACAAGG | 526 | | |
| DB | 122 | CAAAAGAGCAGTTGCAGCAGCTGTCTGACTGGCTGACAGTTACAGAAGAACGCATTTCAG | 181 | | |
| QY | 527 | AAATGGGAAGAGCCTCTTTGGACCTGTATTTGAAGACCTAAACGCCCAAGTACAACAA | 586 | | |

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

3M nucleic - nucleic search, using sw model

Run on: February 1, 2004, 23:51:42 ; Search time 687.863 Seconds
(without alignments)
11131.886 Million cell updates/sec

Title: US-09-845-416-2_COPY_900_3000
Perfect score: 2101
Sequence: 1 gagctatgcctacacacagg.....acctcagcactctggaagac 2101

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2434939 seqs, 1822278265 residues

Total number of hits satisfying chosen parameters: 4869878

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*

10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*

14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*

15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*

16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*

17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|-------------------|
| 1 | 2101 | 100.0 | 4182 | 13 | US-09-845-416-2 |
| 2 | 2101 | 100.0 | 5149 | 13 | US-09-845-416-27 |
| 3 | 1579.8 | 75.2 | 3999 | 13 | US-09-845-416-6 |
| 4 | 1579.8 | 75.2 | 4966 | 13 | US-09-845-416-28 |
| 5 | 1579.8 | 75.2 | 4990 | 13 | US-09-845-416-34 |
| 6 | 1443 | 68.7 | 3858 | 13 | US-09-845-416-9 |
| 7 | 1443 | 68.7 | 4825 | 13 | US-09-845-416-29 |
| 8 | 1443 | 68.7 | 4848 | 13 | US-09-845-416-35 |
| 9 | 1443 | 68.7 | 5060 | 13 | US-09-845-416-36 |
| 10 | 1293 | 61.5 | 8689 | 13 | US-10-149-736-42 |
| 11 | 1114 | 53.0 | 4414 | 13 | US-09-845-416-32 |
| 12 | 1103.4 | 52.5 | 5417 | 13 | US-10-149-736-39 |
| 13 | 1103.4 | 52.5 | 11443 | 13 | US-10-149-736-44 |
| 14 | 1103.4 | 52.5 | 12057 | 13 | US-10-149-736-47 |
| 15 | 1103.4 | 52.5 | 13957 | 10 | US-09-782-378A-22 |

| | | | | | | |
|----|--------|------|-------|----|---------------------|-------------------|
| 16 | 1103.4 | 52.5 | 13957 | 10 | US-09-880-107-2284 | Sequence 2284, Ap |
| 17 | 1103.4 | 52.5 | 13957 | 13 | US-10-149-736-1 | Sequence 1, Appli |
| 18 | 1103.4 | 52.5 | 14082 | 13 | US-10-341-434-108 | Sequence 108, App |
| 19 | 1102 | 52.5 | 3446 | 13 | US-09-845-416-14 | Sequence 14, Appl |
| 20 | 1100.2 | 52.4 | 11058 | 13 | US-09-845-416-1 | Sequence 1, Appli |
| 21 | 1097.8 | 52.3 | 1991 | 13 | US-09-845-416-3 | Sequence 3, Appli |
| 22 | 999 | 47.5 | 2169 | 13 | US-09-845-416-4 | Sequence 4, Appli |
| 23 | 999 | 47.5 | 3531 | 13 | US-09-845-416-10 | Sequence 10, Appl |
| 24 | 999 | 47.5 | 4498 | 13 | US-09-845-416-30 | Sequence 30, Appl |
| 25 | 997 | 47.5 | 5339 | 13 | US-10-149-736-40 | Sequence 40, Appl |
| 26 | 996 | 47.4 | 5462 | 13 | US-10-149-736-41 | Sequence 41, Appl |
| 27 | 866.6 | 41.2 | 13815 | 13 | US-10-149-736-2 | Sequence 2, Appli |
| 28 | 787 | 37.5 | 3510 | 13 | US-09-845-416-12 | Sequence 12, Appl |
| 29 | 787 | 37.5 | 4476 | 13 | US-09-845-416-31 | Sequence 31, Appl |
| 30 | 777 | 37.0 | 1667 | 13 | US-09-845-416-7 | Sequence 7, Appli |
| 31 | 652 | 31.0 | 1821 | 13 | US-09-845-416-13 | Sequence 13, Appl |
| 32 | 450 | 21.4 | 1340 | 13 | US-09-845-416-11 | Sequence 11, Appl |
| 33 | 387 | 18.4 | 387 | 13 | US-10-149-736-32 | Sequence 32, Appl |
| 34 | 361.4 | 17.2 | 10302 | 10 | US-09-782-378A-23 | Sequence 23, Appl |
| 35 | 361.4 | 17.2 | 10302 | 13 | US-10-149-736-3 | Sequence 667, App |
| 36 | 361.4 | 17.2 | 16531 | 13 | US-10-101-510-667 | Sequence 31, Appl |
| 37 | 348 | 16.6 | 348 | 13 | US-10-149-736-31 | Sequence 4, Appli |
| 38 | 336.6 | 16.0 | 11096 | 13 | US-10-149-736-4 | Sequence 9, Appli |
| 39 | 333 | 15.8 | 333 | 13 | US-10-149-736-9 | Sequence 8, Appli |
| 40 | 327 | 15.6 | 327 | 13 | US-10-149-736-8 | Sequence 10, Appl |
| 41 | 322.4 | 15.3 | 333 | 13 | US-10-149-736-10 | Sequence 15, Appl |
| 42 | 265 | 12.6 | 1434 | 13 | US-09-845-416-15 | Sequence 33, Appl |
| 43 | 261 | 12.4 | 324 | 13 | US-10-149-736-33 | Sequence 21956, A |
| 44 | 190.2 | 9.1 | 256 | 9 | US-09-864-761-21956 | Sequence 15766, A |
| 45 | 180 | 8.6 | 476 | 9 | US-09-864-761-15766 | |

ALIGNMENTS

RESULT 1

US-09-845-416-2

; Sequence 2, Application US/09845416

; Publication No. US20030171312A1

; GENERAL INFORMATION:

; APPLICANT: XIAO, XIAO

; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE THEREOF

; FILE REFERENCE: DEL142

; CURRENT APPLICATION NUMBER: US/09/845,416

; CURRENT FILING DATE: 2001-04-30

; PRIOR APPLICATION NUMBER: 60/200,777

; PRIOR FILING DATE: 2000-04-28

; NUMBER OF SEQ ID NOS: 36

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 2

; LENGTH: 4182

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-845-416-2

| | | | | | |
|-----------------------|------|--|--------------|------------|----------------------|
| Query Match | | 100.0%; | Score 2101; | DB 13; | Length 4182; |
| Best Local Similarity | | 100.0%; | Pred. No. 0; | | |
| Matches 2101; | | Conservative | 0; | Mismatches | 0; Indels 0; Gaps 0; |
| QY | 1 | GAGCTATGCCTACACACAGGCTGCTTATGTACACCTCTGACCTACACCGAGCCCAT | 60 | | |
| DB | 900 | GAGCTATGCCTACACACAGGCTGCTTATGTACACCTCTGACCTACACCGAGCCCAT | 959 | | |
| QY | 61 | TCCTTACAGCATTGGAAGCTCCTGAAGACAAAGTCATTGGCAGTTCATTGATGGAGAG | 120 | | |
| DB | 960 | TCCTTACAGCATTGGAAGCTCCTGAAGACAAAGTCATTGGCAGTTCATTGATGGAGAG | 1019 | | |
| QY | 121 | TGAAGTAAACCTGGACCGGTTATCAACACGCTTTAGAAAGATATTATCGTGGCTTCTTTC | 180 | | |
| DB | 1020 | TGAAGTAAACCTGGACCGGTTATCAACACGCTTTAGAAAGATATTATCGTGGCTTCTTTC | 1079 | | |
| QY | 181 | TGCTGAGGACACATTGCAAGCACACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA | 240 | | |

Db 1080 TGCTGAGGACACATTGCAAGCACAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 1139
QY 241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTGTGACAGCCCATCAGGSCCGGTTGG 300
Db 1140 CCAGTTTCATACTCATGAGGGGTACATGATGGATTGTGACAGCCCATCAGGSCCGGTTGG 1199
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1200 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 1259
QY 361 AACTGAAAGTACAAGAGCAGATGAATCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
Db 1260 AACTGAAAGTACAAGAGCAGATGAATCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 1319
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 480
Db 1320 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 1379
QY 481 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAAGAAAAATGGAGGAAGA 540
Db 1380 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAAGAAAAATGGAGGAAGA 1439
QY 541 GCCTCTTGGACCTGATCTTGAAAGACCTAAACCGCAAGTACAACAACATAAGSTGCTTCA 600
Db 1440 GCCTCTTGGACCTGATCTTGAAAGACCTAAACCGCAAGTACAACAACATAAGSTGCTTCA 1499
QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGGTAGT 660
Db 1500 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGGTAGT 1559
QY 661 TGATGAATCTAGTGGAGATCAACGCAACTGCTGCTTTTGGAGAAACAACTTAAGGTATTGGG 720
Db 1560 TGATGAATCTAGTGGAGATCAACGCAACTGCTGCTTTTGGAGAAACAACTTAAGGTATTGGG 1619
QY 721 AGATCGATGGCAAAACATCTGTAGATGGACAGAAAGACCGTGGGTTCTTTTACAGACAT 780
Db 1620 AGATCGATGGCAAAACATCTGTAGATGGACAGAAAGACCGTGGGTTCTTTTACAGACAT 1679
QY 781 CCTTCTCAATGGCAACAGTCTTACTGAAGAACAGTGCCCTTTTAGTGCAATGGCTTTCAGA 840
Db 1680 CCTTCTCAATGGCAACAGTCTTACTGAAGAACAGTGCCCTTTTAGTGCAATGGCTTTCAGA 1739
QY 841 AAAAGAAGATGCAGTGAACAAGATTTCACAACTGGCTTTTAAAGATCAAAATGAATGTT 900
Db 1740 AAAAGAAGATGCAGTGAACAAGATTTCACAACTGGCTTTTAAAGATCAAAATGAATGTT 1799
QY 901 ATCAAGTCTTCAAAAACTGGCCGTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 960
Db 1800 ATCAAGTCTTCAAAAACTGGCCGTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 1859
QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTTCAACACTGAAGAAATAAGTCAGTGAC 1020
Db 1860 GGGCAAACTGTATTCACTCAAAACAAGATCTTTCAACACTGAAGAAATAAGTCAGTGAC 1919
QY 1021 CCAGAAGACGGAGCATGGCTGGATAACTTTGCCGGTGTGGGATAATTTAGTCCAAA 1080
Db 1920 CCAGAAGACGGAGCATGGCTGGATAACTTTGCCGGTGTGGGATAATTTAGTCCAAA 1979
QY 1081 ACTTGAAAAGAGTACGACACAGACTCATAGATTACTGCAACAGTTCCCTCCCTGGACCTGGA 1140
Db 1980 ACTTGAAAAGAGTACGACACAGACTCATAGATTACTGCAACAGTTCCCTCCCTGGACCTGGA 2039
QY 1141 AAAGTTTCTTGGCTTACAGAAGCTGAACAACTGCCAATGTCTACAGGATGCTAC 1200
Db 2040 AAAGTTTCTTGGCTTACAGAAGCTGAACAACTGCCAATGTCTACAGGATGCTAC 2099
QY 1201 CCGTAAGGAAAGGCTCTAGAAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 1260
Db 2100 CCGTAAGGAAAGGCTCTAGAAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 2159
QY 1261 AGACCTCCAAGGTGAATTAAGCTCACACAGATGTTTATCACAACTGGATGAAACAG 1320

Db 2160 AGACCTCCAAGGTGAATAATTGAAGCTCACACAGATGTTTATCAAACTGGATGAAAACAG 2219
QY 1321 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTACAAAAGACGTTT 1380
Db 2220 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTACAAAAGACGTTT 2279
QY 1381 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCCA 1440
Db 2280 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCCA 2339
QY 1441 TTTGGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGT 1500
Db 2340 TTTGGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGT 2399
QY 1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTTCC 1560
Db 2400 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTTCC 2459
QY 1561 AGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGGAATTGAAAACTAAAGA 1620
Db 2460 AGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGGAATTGAAAACTAAAGA 2519
QY 1621 ACCGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGA 1680
Db 2520 ACCGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGA 2579
QY 1681 AGGACTAGAGAAAACTCTACAGGAGCCCCAGAGAGCTGCCTCTCTGAGGAGAGAGGCCAGAA 1740
Db 2580 AGGACTAGAGAAAACTCTACAGGAGCCCCAGAGAGCTGCCTCTCTGAGGAGAGAGGCCAGAA 2639
QY 1741 TGTCACTCGGCTTCTACGAAACAGCGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAA 1800
Db 2640 TGTCACTCGGCTTCTACGAAACAGCGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAA 2699
QY 1801 CCTGCACTCCGCTGACTGSCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 1860
Db 2700 CCTGCACTCCGCTGACTGSCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 2759
QY 1861 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGGATC 1920
Db 2760 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGGATC 2819
QY 1921 CTGGCAGCCCGTGGCGGATCTCTCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAA 1980
Db 2820 CTGGCAGCCCGTGGCGGATCTCTCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAA 2879
QY 1981 GGCACCTCGAGGAGAAAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCACCTTGC 2040
Db 2880 GGCACCTCGAGGAGAAAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCACCTTGC 2939
QY 2041 TCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 2100
Db 2940 TCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 2999
QY 2101 C 2101
Db 3000 C 3000

RESULT 2
US-09-845-416-27
; Sequence 27, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36

| | | | | | | | | | |
|---|------|---|------|--|--|--|--|--|--|
| SOFTWARE: PatentIn Ver. 2.1 | | | | | | | | | |
| SEQ ID NO 27 | | | | | | | | | |
| LENGTH: 5149 | | | | | | | | | |
| TYPE: DNA | | | | | | | | | |
| ORGANISM: Homo sapiens | | | | | | | | | |
| IS-09-845-416-27 | | | | | | | | | |
| Query Match 100.0%; Score 2101; DB 13; Length 5149; | | | | | | | | | |
| Best Local Similarity 100.0%; Pred. No. 0; | | | | | | | | | |
| Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | | | | | | | |
| 2Y | 1 | GAGCTATGCCTACACACAGGCTGCTTATGTCCACCTCTGACCCCTACACGGAGCCCCATT | 50 | | | | | | |
| 2b | 1657 | GAGCTATGCCTACACACAGGCTGCTTATGTCCACCTCTGACCCCTACACGGAGCCCCATT | 1716 | | | | | | |
| 2Y | 61 | TCCTTCACAGCATTTGGAAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG | 120 | | | | | | |
| 2b | 1717 | TCCTTCACAGCATTTGGAAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG | 1776 | | | | | | |
| 2Y | 121 | TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAAAGTATTATCGTGGCTTCTTTTC | 180 | | | | | | |
| 2b | 1777 | TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAAAGTATTATCGTGGCTTCTTTTC | 1836 | | | | | | |
| 2Y | 181 | TGCTGAGGACACATTCGAAGCACAAAGGAGAGATTTCCTAATGATGTGGAAGTGGTGAAGA | 240 | | | | | | |
| 2b | 1837 | TGCTGAGGACACATTCGAAGCACAAAGGAGAGATTTCCTAATGATGTGGAAGTGGTGAAGA | 1896 | | | | | | |
| 2Y | 241 | CCAGTTTCATACTCATGAGGGTACATGATGGAATTTGACAGCCCCATCAGGGCCGGTTGG | 300 | | | | | | |
| 2b | 1897 | CCAGTTTCATACTCATGAGGGTACATGATGGAATTTGACAGCCCCATCAGGGCCGGTTGG | 1956 | | | | | | |
| 2Y | 301 | TAATATTCTACAATTGGGAAGTAAAGCTGATTTGGAACAGGAAAAATTATCAGAAGATGAAGA | 360 | | | | | | |
| 2b | 1957 | TAATATTCTACAATTGGGAAGTAAAGCTGATTTGGAACAGGAAAAATTATCAGAAGATGAAGA | 2016 | | | | | | |
| 2Y | 361 | AACCTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC | 420 | | | | | | |
| 2b | 2017 | AACCTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC | 2076 | | | | | | |
| 2Y | 421 | TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT | 480 | | | | | | |
| Db | 2077 | TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT | 2136 | | | | | | |
| 2Y | 481 | GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAACCAAGGAAAAATGGAGGAAGA | 540 | | | | | | |
| Db | 2137 | GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAACCAAGGAAAAATGGAGGAAGA | 2196 | | | | | | |
| 2Y | 541 | GCCTCTTGACCTGATCTTGAAGACCTAAAACGCCAAGTACAAACAATAGGTGCTTCA | 600 | | | | | | |
| Db | 2197 | GCCTCTTGACCTGATCTTGAAGACCTAAAACGCCAAGTACAAACAATAGGTGCTTCA | 2256 | | | | | | |
| 2Y | 601 | AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGGTAGT | 660 | | | | | | |
| 2b | 2257 | AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGGTAGT | 2316 | | | | | | |
| 2Y | 661 | TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAGAAACAACTTAAGGTATTGGG | 720 | | | | | | |
| 2b | 2317 | TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAGAAACAACTTAAGGTATTGGG | 2376 | | | | | | |
| 2Y | 721 | AGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT | 780 | | | | | | |
| Db | 2377 | AGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT | 2436 | | | | | | |
| 2Y | 781 | CCTTCTCAATGGCAACGCTTTACTGAAGAACAGTCGCTTTTATGTCATGGCTTTTCAGA | 840 | | | | | | |
| Db | 2437 | CCTTCTCAATGGCAACGCTTTACTGAAGAACAGTCGCTTTTATGTCATGGCTTTTCAGA | 2496 | | | | | | |
| 2Y | 841 | AAAAGAAGATGCAGTGAAACAAGATTACACAACCTGGCTTTAAAGATCAAAATGAAATGTT | 900 | | | | | | |
| Db | 2497 | AAAAGAAGATGCAGTGAAACAAGATTACACAACCTGGCTTTAAAGATCAAAATGAAATGTT | 2556 | | | | | | |
| 2Y | 901 | ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT | 960 | | | | | | |

| | | | |
|----|------|--|------|
| Db | 2557 | ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT | 2616 |
| QY | 961 | GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC | 1020 |
| Db | 2617 | GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC | 2676 |
| QY | 1021 | CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAAA | 1080 |
| Db | 2677 | CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAAA | 2736 |
| QY | 1081 | ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCCTGGACCTTGA | 1140 |
| Db | 2737 | ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCCTGGACCTTGA | 2796 |
| QY | 1141 | AAAGTTTCTTGCTGGCTTACAGAAAGCTGAAACAACTGCCAATGTCTACAGGATGCTAC | 1200 |
| Db | 2797 | AAAGTTTCTTGCTGGCTTACAGAAAGCTGAAACAACTGCCAATGTCTACAGGATGCTAC | 2856 |
| QY | 1201 | CCGTAAGGAAAGGCTCCTTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCA | 1260 |
| Db | 2857 | CCGTAAGGAAAGGCTCCTTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCA | 2916 |
| QY | 1261 | AGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACAG | 1320 |
| Db | 2917 | AGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACAG | 2976 |
| QY | 1321 | CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGTCAGTCTCTGTTACAAAGACGTTT | 1380 |
| Db | 2977 | CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGTCAGTCTCTGTTACAAAGACGTTT | 3036 |
| QY | 1381 | GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAATAAAGTCTCTCAACATTAGGTCCCA | 1440 |
| Db | 3037 | GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAATAAAGTCTCTCAACATTAGGTCCCA | 3096 |
| QY | 1441 | TTTGGAAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGT | 1500 |
| Db | 3097 | TTTGGAAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGT | 3156 |
| QY | 1501 | GTGGTACAGCTGAAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC | 1560 |
| Db | 3157 | GTGGTACAGCTGAAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC | 3216 |
| QY | 1561 | AGCAGTTCAGAAAGCAGAACCATGTACATAGGGCCCTTCAAGAGGGGAATTTGAAAACTAAAGA | 1620 |
| Db | 3217 | AGCAGTTCAGAAAGCAGAACCATGTACATAGGGCCCTTCAAGAGGGGAATTTGAAAACTAAAGA | 3276 |
| QY | 1621 | ACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGA | 1680 |
| Db | 3277 | ACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGA | 3336 |
| QY | 1681 | AGGACTAGAGAAAACTCTACAGAGGCCAGAGAGCTGCGCTCCTGAGGAGAGAGCCCCAGAA | 1740 |
| Db | 3337 | AGGACTAGAGAAAACTCTACAGAGGCCAGAGAGCTGCGCTCCTGAGGAGAGAGCCCCAGAA | 3396 |
| QY | 1741 | TGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA | 1800 |
| Db | 3397 | TGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA | 3456 |
| QY | 1801 | CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAGACTCCAGGAACT | 1860 |
| Db | 3457 | CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAGACTCCAGGAACT | 3516 |
| QY | 1861 | TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAGCTGAGGTGATCAAGGGATC | 1920 |
| Db | 3517 | TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAGCTGAGGTGATCAAGGGATC | 3576 |
| QY | 1921 | CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA | 1980 |
| Db | 3577 | CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA | 3636 |
| QY | 1981 | GGCATTTCAGGAGAAAAATTGGCGCTCTGAAAGAGAACGTCAGGCCACGTCAATGACCTTGC | 2040 |
| Db | 3637 | GGCATTTCAGGAGAAAAATTGGCGCTCTGAAAGAGAACGTCAGGCCACGTCAATGACCTTGC | 3696 |

| | | | |
|--|------|--|------|
| QY | 2041 | TCGCCAGCTTACCACTTTGGGCATTTCAGCTCTCACCGTATATAACCTCAGCACTCTCGGAAGA | 2100 |
| Db | 3697 | TCGCCAGCTTACCACTTTGGGCATTTCAGCTCTCACCGTATATAACCTCAGCACTCTCGGAAGA | 3756 |
| QY | 2101 | C | 2101 |
| Db | 3757 | C | 3757 |
| RESULT 3 | | | |
| US-09-845-416-6 | | | |
| ; Sequence 6, Application US/09845416 | | | |
| ; Publication No. US20030171312A1 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: XIAO, XIAO | | | |
| ; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE | | | |
| ; FILE REFERENCE: DE1142 | | | |
| ; CURRENT APPLICATION NUMBER: US/09/845,416 | | | |
| ; CURRENT FILING DATE: 2001-04-30 | | | |
| ; PRIOR APPLICATION NUMBER: 60/200,777 | | | |
| ; PRIOR FILING DATE: 2000-04-28 | | | |
| ; NUMBER OF SEQ ID NOS: 36 | | | |
| ; SOFTWARE: PatentIn Ver. 2.1 | | | |
| ; SEQ ID NO 6 | | | |
| ; LENGTH: 3999 | | | |
| ; TYPE: DNA | | | |
| ; ORGANISM: Homo sapiens | | | |
| US-09-845-416-6 | | | |
| Query Match 75.2%; Score 1579.8; DB 13; Length 3999; | | | |
| Best Local Similarity 87.9%; Pred. No. 0; | | | |
| Matches 1846; Conservative 0; Mismatches 72; Indels 183; Gaps 4; | | | |
| QY | 1 | GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAATT | 60 |
| Db | 900 | GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAATT | 959 |
| QY | 61 | TCCTTACACAGCATTTGGAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG | 120 |
| Db | 960 | TCCTTACACAGCATTTGGAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG | 1019 |
| QY | 121 | TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATATTATCGTGCTTCTTTT | 180 |
| Db | 1020 | TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATATTATCGTGCTTCTTTT | 1079 |
| QY | 181 | TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA | 240 |
| Db | 1080 | TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA | 1139 |
| QY | 241 | CCAGTTTCTACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG | 300 |
| Db | 1140 | CCAGTTTCTACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG | 1199 |
| QY | 301 | TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA | 360 |
| Db | 1200 | TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA | 1259 |
| QY | 361 | AACTGAAGTACAAGACAGATGAATCTCTAAATTCAGATGGGAATCCCTCAGGTTAGC | 420 |
| Db | 1260 | AACTGAAGTACAAGACAGATGAATCTCTAAATTCAGATGGGAATCCCTCAGGTTAGC | 1319 |
| QY | 421 | TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT | 480 |
| Db | 1320 | TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT | 1379 |
| QY | 481 | GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAGGAAATGGAGGAAGA | 540 |
| Db | 1380 | GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAGGAAATGGAGGAAGA | 1439 |
| QY | 541 | GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAACATAAGGTGCTTCA | 600 |
| Db | | | |
| Db | 1440 | GCCTCTTGGACCTGATCTTTGAAGACCTTAAAAAGCCCAAGTACAACAATAAGGTGCTTCA | 1499 |
| QY | 601 | AGAAGATCTAGAAACAAGACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGTAGT | 660 |
| Db | 1500 | AGAAGATCTAGAAACAAGACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGTAGT | 1559 |
| QY | 661 | TGATGAATCTAGTGGAGATCACGCAACTGCTTTTGGAGAACAACTTAAGGTATTGGG | 720 |
| Db | 1560 | TGATGAATCTAGTGGAGATCACGCAACTGCTTTTGGAGAACAACTTAAGGTATTGGG | 1619 |
| QY | 721 | AGATCGATGGCAAAACATCTGTAGATGGACAGAAACCGCTGGGTTCTTTTACAAGACAT | 780 |
| Db | 1620 | AGATCGATGGCAAAACATCTGTAGATGGACAGAAACCGCTGGGTTCTTTTACAAGACCA | 1679 |
| QY | 781 | CCTTCTCAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTACA | 840 |
| Db | 1680 | GCCTGAC-----CTAGCTCCTGGACTGACCACATATTGGAGCCTCTCCTACTCA | 1727 |
| QY | 841 | AAAAGAAGATGCAGTGAACAAGATTACACAAACTGGCTTTTAAAGATCAAAATGAAATGTT | 900 |
| Db | 1728 | GACTGTTACTCTGTGACACA-----ACCTGTGGTTACTAAGGAAACTGCCATCT- | 1777 |
| QY | 901 | ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT | 960 |
| Db | 1778 | ----- | 1777 |
| QY | 961 | GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACCTGAAGAATAAGTCAGTGAC | 1020 |
| Db | 1778 | -----CCAAACTAGAAATGCCATCTTCTTGTGTTGGAG----- | 1812 |
| QY | 1021 | CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAAATTTAGTCCAAAA | 1080 |
| Db | 1813 | ----- | 1812 |
| QY | 1081 | ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA | 1140 |
| Db | 1813 | -----GTACCTACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA | 1856 |
| QY | 1141 | AAAGTTTCTTGCTGCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTAC | 1200 |
| Db | 1857 | AAAGTTTCTTGCTGCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTAC | 1916 |
| QY | 1201 | CCGTAAGGAAAGGCTCCTAGAAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA | 1260 |
| Db | 1917 | CCGTAAGGAAAGGCTCCTAGAAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA | 1976 |
| QY | 1261 | AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAG | 1320 |
| Db | 1977 | AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAG | 2036 |
| QY | 1321 | CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCGAGTCTCTCAACATTAGGTCCCA | 1380 |
| Db | 2037 | CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCGAGTCTCTCAACATTAGGTCCCA | 2096 |
| QY | 1381 | GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCA | 1440 |
| Db | 2097 | GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCA | 2156 |
| QY | 1441 | TTTGGAAAGCCAGTTCTGACCAAGTGAAGCGTCTGCAACCTTTCTCTGCAGGAACCTTCTGGT | 1500 |
| Db | 2157 | TTTGGAAAGCCAGTTCTGACCAAGTGAAGCGTCTGCAACCTTTCTCTGCAGGAACCTTCTGGT | 2216 |
| QY | 1501 | GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGGCGACTTCC | 1560 |
| Db | 2217 | GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGGCGACTTCC | 2276 |
| QY | 1561 | AGCAGTTCAAGACAGAACCGATGATGATAGGGCTTCAAGAGGGAATTGAAAACTAAAGA | 1620 |
| Db | 2277 | AGCAGTTCAAGACAGAACCGATGATGATAGGGCTTCAAGAGGGAATTGAAAACTAAAGA | 2336 |
| QY | 1621 | ACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGA | 1680 |
| Db | 2337 | ACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGA | 2396 |

```
1681 AGGACTAGAGAAACTCTACAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAA 1740
|||||
2397 AGGACTAGAGAAACTCTACAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAA 2456
|||||
1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAA 1800
|||||
2457 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAA 2516
|||||
1801 CTTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 1860
|||||
2517 CTTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 2576
|||||
1861 TCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCCCAAGCTGAGGTGATCAAGGGATC 1920
|||||
2577 TCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCCCAAGCTGAGGTGATCAAGGGATC 2636
|||||
1921 CTGGCAGCCCGTGGCGATCTCCTCATTTGAGTCTCTCCCAAGATCACCTCGAGAAAAAGTCAA 1980
|||||
2637 CTGGCAGCCCGTGGCGATCTCCTCATTTGAGTCTCTCCCAAGATCACCTCGAGAAAAAGTCAA 2696
|||||
1981 GGCACTTCGAGGAGAAATTGGCGCTCTGAAAGAGAAAGCTGAGCCACGTCATGACCTTGC 2040
|||||
2697 GGCACTTCGAGGAGAAATTGGCGCTCTGAAAGAGAAAGCTGAGCCACGTCATGACCTTGC 2756
|||||
2041 TCGCCAGCTTACCACCTTTGGSCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 2100
|||||
2757 TCGCCAGCTTACCACCTTTGGSCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 2816
|||||
2101 C 2101
2817 C 2817

RESULT 4
US-09-845-416-28
; Sequence 28, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 4966
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-28

Query Match 75.2%; Score 1579.8; DB 13; Length 4966;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 1846; Conservative 0; Mismatches 72; Indels 183; Gaps 4;

2Y 1 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGGAGCCCAATT 60
|||||
Db 1657 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGGAGCCCAATT 1716
|||||
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAGACAAAGTCAATTTGGCAGTTCATTTGATGGAGAG 120
|||||
Db 1717 TCCTTCACAGCATTTGGAAGCTCCTGAGACAAAGTCAATTTGGCAGTTCATTTGATGGAGAG 1776
|||||
QY 121 TGAAGTAAACCTGACCCGTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCCTTC 180
|||||
Db 1777 TGAAGTAAACCTGACCCGTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCCTTC 1836
|||||
QY 181 TGCTGAGGACACATTGCAAGCAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 240
|||||
```

```
Db 1837 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTCAAAGA 1896
QY 241 CCAGTTTTCATACCTCATGAGGGGTACATGATGATGATTTGACAGCCCCATCAGGSCCGGTGG 300
|||||
Db 1897 CCAGTTTTCATACCTCATGAGGGGTACATGATGATGATTTGACAGCCCCATCAGGSCCGGTGG 1956
|||||
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTTATCAGAAGATGAAGA 360
|||||
Db 1957 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTTATCAGAAGATGAAGA 2016
|||||
QY 361 AACTGAAAGTACAAGAGCAGATGAATCTCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
|||||
Db 2017 AACTGAAAGTACAAGAGCAGATGAATCTCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2076
|||||
QY 421 TAGCATGGAAAAACAAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAAACT 480
|||||
Db 2077 TAGCATGGAAAAACAAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAAACT 2136
|||||
QY 481 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAAGGAAAAATGGAGGAAGA 540
|||||
Db 2137 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAAGGAAAAATGGAGGAAGA 2196
|||||
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCAAAGTACAACACATAAGGTGCTTCA 600
|||||
Db 2197 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCAAAGTACAACACATAAGGTGCTTCA 2256
|||||
QY 601 AGAAGATCTAGAAACAAGAACAAAGTCAAGGTCAATCTCTCACTCACATGGTGGGTAGT 660
|||||
Db 2257 AGAAGATCTAGAAACAAGAACAAAGTCAAGGTCAATCTCTCACTCACATGGTGGGTAGT 2316
|||||
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 720
|||||
Db 2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 2376
|||||
QY 721 AGATCGATGGGCACAAACATCTGTAGATGGACAGACCGCTGGTTCTTTTACAAGACAT 780
|||||
Db 2377 AGATCGATGGGCACAAACATCTGTAGATGGACAGACCGCTGGTTCTTTTACAAGACCA 2436
|||||
QY 781 CCTTCTCAAAATGGCAACGTTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTTCA 840
|||||
Db 2437 GCCTGAC-----CTAGCTCCTGGACTGACCACTATTGGAGCCTCTCCTACTCA 2484
|||||
QY 841 AAAAGAGATGCAGTGAACAAGATTCACACAACTGGCTTTTAAAGATCAAAATGAATGTT 900
|||||
Db 2485 GACTGTACTCTGGTGACACA-----ACCTGTGGTTACTAAGGAAACTGCCATCT- 2534
|||||
QY 901 ATCAAGTCTTCAAAAACTGGCCGCTTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT 960
|||||
Db 2535 ----- 2534
|||||
QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTTCAACACTGAAGAATAAGTCAGTGAC 1020
|||||
Db 2535 -----CCAAACTAGAAATGCCATCTTCTTGATGTTGGAG----- 2569
|||||
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAAATTTAGTCCAAAA 1080
|||||
Db 2570 ----- 2569
|||||
QY 1081 ACTTGAAGAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA 1140
|||||
Db 2570 -----GTACCTACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA 2613
|||||
QY 1141 AAAGTTTCTTTCCTGGCTTACAGAAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTAC 1200
|||||
Db 2614 AAAGTTTCTTTCCTGGCTTACAGAAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTAC 2673
|||||
QY 1201 CCGTAAGGAAAGGCTCCTAGAAAGCTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 1260
|||||
Db 2674 CCGTAAGGAAAGGCTCCTAGAAAGCTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 2733
|||||
QY 1261 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAG 1320
|||||
Db 2734 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAG 2793
|||||
```

| | | | |
|----|------|--|------|
| QY | 1321 | CCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTT | 1380 |
| DB | 2794 | CCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTT | 2853 |
| QY | 1381 | GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCA | 1440 |
| DB | 2854 | GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCA | 2913 |
| QY | 1441 | TTTGGAGCCAGTTCTTGACCAGTGGAGCGGTCTGCACCTTTCTCTGCAGGAACCTTCTGGT | 1500 |
| DB | 2914 | TTTGGAGCCAGTTCTTGACCAGTGGAGCGGTCTGCACCTTTCTCTGCAGGAACCTTCTGGT | 2973 |
| QY | 1501 | GTGGCTACAGCTGAAAGATGATGAATTAAGCGCGGCAGGCACCTATTGGAGGCGACTTTCC | 1560 |
| DB | 2974 | GTGGCTACAGCTGAAAGATGATGAATTAAGCGCGGCAGGCACCTATTGGAGGCGACTTTCC | 3033 |
| QY | 1561 | AGCAGTTCAGAAAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAAAACTAAAGA | 1620 |
| DB | 3034 | AGCAGTTCAGAAAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAAAACTAAAGA | 3093 |
| QY | 1621 | ACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGA | 1680 |
| DB | 3094 | ACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGA | 3153 |
| QY | 1681 | AGGACTAGAGAAACTCTACCAAGGAGCCAGAGAGTGCCTCCTGTAGGAGAGAGCCAGAA | 1740 |
| DB | 3154 | AGGACTAGAGAAACTCTACCAAGGAGCCAGAGAGTGCCTCCTGTAGGAGAGAGCCAGAA | 3213 |
| QY | 1741 | TGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAA | 1800 |
| DB | 3214 | TGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAA | 3273 |
| QY | 1801 | CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT | 1860 |
| DB | 3274 | CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT | 3333 |
| QY | 1861 | TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGGATC | 1920 |
| DB | 3334 | TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGGATC | 3393 |
| QY | 1921 | CTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA | 1980 |
| DB | 3394 | CTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA | 3453 |
| QY | 1981 | GGCACTTCGAGGAGAAAATTGGCGCTCTGAAAGAGAACGTGAGGCCACGTCAATGACCTTGC | 2040 |
| DB | 3454 | GGCACTTCGAGGAGAAAATTGGCGCTCTGAAAGAGAACGTGAGGCCACGTCAATGACCTTGC | 3513 |
| QY | 2041 | TCGCCAGCTTACCATTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA | 2100 |
| DB | 3514 | TCGCCAGCTTACCATTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA | 3573 |
| QY | 2101 | C 2101 | |
| DB | 3574 | C 3574 | |

RESULT 5

```

US-09-845-416-34
; Sequence 34, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1

```

```

; SEQ ID NO 34
; LENGTH: 4990
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-845-416-34

      75.2%;   Score 1579.8;   DB 13;   Length 4990;
Query Match      87.9%;   Pred. No. 0;
Best Local Similarity
Matches 1846;   Conservative   0;   Mismatches   72;   Indels   183;   Gaps   4;

      1  GAGCTATGCCTACACACAGGCTGCTTATGTACCACTCTGACCTACACGGAGCCCAATT 60
      Db  GAGCTATGCCTACACACAGGCTGCTTATGTACCACTCTGACCTACACGGAGCCCAATT 1740
      61  TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 120
      Qy  TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1800
      Db  TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1800
      121  TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
      Qy  TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1860
      Db  TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1860
      181  TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
      Qy  TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
      Db  TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1920
      241  CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
      Qy  CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
      Db  CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1980
      301  TAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 360
      Qy  TAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 360
      Db  TAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 2040
      361  AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGSGAATGCCTCAGGGTAGC 420
      Qy  AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGSGAATGCCTCAGGGTAGC 420
      Db  AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGSGAATGCCTCAGGGTAGC 2100
      421  TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAAATCAGAAAAT 480
      Qy  TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAAATCAGAAAAT 480
      Db  TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAAATCAGAAAAT 2160
      481  GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAACAAAGGAAATGGAGGAAGA 540
      Qy  GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAACAAAGGAAATGGAGGAAGA 540
      Db  GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAACAAAGGAAATGGAGGAAGA 2220
      541  GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAAACAATAAGGTGCTTCA 600
      Qy  GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAAACAATAAGGTGCTTCA 600
      Db  GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAAACAATAAGGTGCTTCA 2280
      601  AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGTAGT 660
      Qy  AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGTAGT 660
      Db  AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGTAGT 2340
      661  TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
      Qy  TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
      Db  TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 2400
      721  AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGCTGGGTCTCTTTTACAAGACAT 780
      Qy  AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGCTGGGTCTCTTTTACAAGACAT 780
      Db  AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGCTGGGTCTCTTTTACAAGACCA 2460
      781  CCTTCTCAAATGGCAACGTTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 840
      Qy  CCTTCTCAAATGGCAACGTTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 840
      Db  CCTTCTCAAATGGCAACGTTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 2508
      841  AAAAGAAGATGCAGTGAACAAGATTACACAACCTGGCTTTAAAGATCAAAATGAAATGTT 900
      Qy  AAAAGAAGATGCAGTGAACAAGATTACACAACCTGGCTTTAAAGATCAAAATGAAATGTT 900
      Db  GACTGTTACTCTGGTGAACA-----ACCTGTGGTACTAAGGAAACCTGCCATCT- 2558
      901  ATCAAGTCTTCAAAAACCTGGCGCTTTTAAAGCGGATCTAGAAAAAGAAAAGCAATCCAT 960
      Qy  ATCAAGTCTTCAAAAACCTGGCGCTTTTAAAGCGGATCTAGAAAAAGAAAAGCAATCCAT 960
      Db  ----- 2558

```


Y 961 GGGCAAACTGFAATTCACCTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
b 2559 -----CCAAACTAGAAATGCCATCTTCTCTTGATGTTGGAG----- 2593
Y 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGTGTGGGATAAATTTAGTCCAAAA 1080
b 2594 ----- 2593
Y 1081 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA 1140
b 2594 -----GTACCTACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA 2637
Y 1141 AAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACACTGCCAATGTCCTACAGGATGCTAC 1200
b 2638 AAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACACTGCCAATGTCCTACAGGATGCTAC 2697
Y 1201 CCGTAAGGAAAGGCTCCTAGAGACTCCAAAGGAGTAAAGAGCTGATGAAACAATGGCA 1260
b 2698 CCGTAAGGAAAGGCTCCTAGAGACTCCAAAGGAGTAAAGAGCTGATGAAACAATGGCA 2757
Y 1261 AGACCTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAACAAG 1320
b 2758 AGACCTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAACAAG 2817
Y 1321 CCAAAAAATCCTGAGATCCCTGGAGGTTCCGATGATGCAGTCTCTGTACAAAGACGTTT 1380
b 2818 CCAAAAAATCCTGAGATCCCTGGAGGTTCCGATGATGCAGTCTCTGTACAAAGACGTTT 2877
Y 1381 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGTCCCA 1440
b 2878 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGTCCCA 2937
Y 1441 TTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTCTGCGAGAACTTCTGGT 1500
b 2938 TTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTCTGCGAGAACTTCTGGT 2997
Y 1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGCGGACTTTCC 1560
b 2998 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGCGGACTTTCC 3057
Y 1561 AGCAGTTCAGAAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAAATTGAAAACTAAAGA 1620
b 3058 AGCAGTTCAGAAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAAATTGAAAACTAAAGA 3117
Y 1621 ACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCCTTTGGA 1680
b 3118 ACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCCTTTGGA 3177
Y 1681 AGGACTAGAGAAACTCTACCAGAGCCAGAGAGCTGCCTCCTGAGGAGAGGCCAGAA 1740
b 3178 AGGACTAGAGAAACTCTACCAGAGCCAGAGAGCTGCCTCCTGAGGAGAGGCCAGAA 3237
Y 1741 TGTCACTCGGCTTCTACGAAAGCAGAGGCTGAGGAGGTCAATACTGAGTGGGAAATAATTGAA 1800
b 3238 TGTCACTCGGCTTCTACGAAAGCAGAGGCTGAGGAGGTCAATACTGAGTGGGAAATAATTGAA 3297
Y 1801 CCTGCACCTCGGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAAGACTCCAGGAACT 1860
b 3298 CCTGCACCTCGGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAAGACTCCAGGAACT 3357
Y 1861 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAGCTGAGGTGATCAAGGGATC 1920
b 3358 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAGCTGAGGTGATCAAGGGATC 3417
Y 1921 CTGGCAGCCCTGGGCGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAA 1980
b 3418 CTGGCAGCCCTGGGCGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAA 3477
Y 1981 GGCACCTTCGAGGAGAAATTTGGCCTCTGAAAGAGAACGTTGAGCCACGTTCAATGACCTTGC 2040
b 3478 GGCACCTTCGAGGAGAAATTTGGCCTCTGAAAGAGAACGTTGAGCCACGTTCAATGACCTTGC 3537

QY 2041 TCGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACCTCTGGAAGA 2100
Db 3538 TCGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACCTCTGGAAGA 3597
QY 2101 C 2101
Db 3598 C 3598
RESULT 6
US-09-845-416-9
; Sequence 9, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3858
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-9
Query Match 68.7%; Score 1443; DB 13; Length 3858;
Best Local Similarity 84.6%; Pred. No. 0;
Matches 1777; Conservative 0; Mismatches 0; Indels 324; Gaps 1;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGAGCCCCATT 60
Db 900 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGAGCCCCATT 959
QY 61 TCCTTTCACAGCATTTTGGAAAGCTCCTGAAGACAACTCATTGGCAGTTTCATTGATGGAGAG 120
Db 960 TCCTTTCACAGCATTTTGGAAAGCTCCTGAAGACAACTCATTGGCAGTTTCATTGATGGAGAG 1019
QY 121 TGAAGTAAACCTGGACCGCTTATCAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTC 180
Db 1020 TGAAGTAAACCTGGACCGCTTATCAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTC 1079
QY 181 TGCTGAGGACACATTGCAAGCACCAAGGAGAGATTTCATAATGATGTGGAAGTGGTGAAGA 240
Db 1080 TGCTGAGGACACATTGCAAGCACCAAGGAGAGATTTCATAATGATGTGGAAGTGGTGAAGA 1139
QY 241 CCAGTTTCATACCTCATGAGGGGTACATGATGGATTTCAGACCCCATCAGGGCCGGTGG 300
Db 1140 CCAGTTTCATACCTCATGAGGGGTACATGATGGATTTCAGACCCCATCAGGGCCGGTGG 1199
QY 301 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAAATTATCAGAAGATGAAGA 360
Db 1200 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAAATTATCAGAAGATGAAGA 1259
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGCTAGC 420
Db 1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGCTAGC 1319
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAACT 480
Db 1320 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAACT 1379
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAACCAAGGAAAATGGAGGAGA 540
Db 1380 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAACCAAGGAAAATGGAGGAGA 1439
QY 541 GCCTCTTTGGACCTGATCTTTGAAGACCTAAAAACGCCCAAGTACAAACATAAGGTGCTTCA 600
Db 1440 GCCTCTTTGGACCTGATCTTTGAAGACCTAAAAACGCCCAAGTACAAACATAAGGTGCTTCA 1499

```
QY 601 AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTTCTCACTCACAATGGTGGTAGT 660
Db 1500 AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTTCTCACTCACAATGGTGGTAGT 1559
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGGAAGAACAACTTAAGGTATTGGG 720
Db 1560 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGGAAGAACAACTTAAGGTATTGGG 1619
QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTCTTTTACAAGACAT 780
Db 1620 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTCTTTTACAAGAC-- 1677
QY 781 CCTTCTCAATGGCAACGCTTACTGAAGAACAGTGCCTTTTGTAGTCAATGGCTTTTACA 840
Db 1678 ----- 1677
QY 841 AAAAGAAGATGCAGTGAACAAGATTACACAACCTGGCTTTAAAGATCAAAAATGAAATGTT 900
Db 1678 ----- 1677
QY 901 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAAGAAAAGCAATCCAT 960
Db 1678 ----- 1677
QY 961 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCACTGAC 1020
Db 1678 ----- 1677
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTGGGATAATTTAGTCCAAAA 1080
Db 1678 ----- 1677
QY 1081 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA 1140
Db 1678 -----ACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA 1715
QY 1141 AAAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGTAC 1200
Db 1716 AAAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGTAC 1775
QY 1201 CCGTAAGGAAAGGCTCCTAGAAGACTCCAGGGAGTAAAGAGCTGATGAAACAATGGCA 1260
Db 1776 CCGTAAGGAAAGGCTCCTAGAAGACTCCAGGGAGTAAAGAGCTGATGAAACAATGGCA 1835
QY 1261 AGACCTCCAAGGTGAATTAAGCTCACACAGATGTTTATCAACAACCTGGATGAAACAAG 1320
Db 1836 AGACCTCCAAGGTGAATTAAGCTCACACAGATGTTTATCAACAACCTGGATGAAACAAG 1895
QY 1321 CCAAAAAATCCTGAGATCCTCGAAGGTTCCGATGATGCAGTCTGTACAAAGACGTTT 1380
Db 1896 CCAAAAAATCCTGAGATCCTCGAAGGTTCCGATGATGCAGTCTGTACAAAGACGTTT 1955
QY 1381 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGTCCCA 1440
Db 1956 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGTCCCA 2015
QY 1441 TTTGGAAGCCAGTTCTGACCACTGGAAGCGTCTGCACCTTCTCTGCAAGAACTTCTGGT 1500
Db 2016 TTTGGAAGCCAGTTCTGACCACTGGAAGCGTCTGCACCTTCTCTGCAAGAACTTCTGGT 2075
QY 1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC 1560
Db 2076 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC 2135
QY 1561 AGCAGTTTCAAGAACGACGATGTACATAGGGCTTCAAGAGGGGAATTGAAAACTAAAGA 1620
Db 2136 AGCAGTTTCAAGAACGACGATGTACATAGGGCTTCAAGAGGGGAATTGAAAACTAAAGA 2195
QY 1621 ACCTGTAATCATGAGTACTCTTGTAGACTGTACGAATATTTTGTACAGAGCAGCCTTTGGA 1680
Db 2196 ACCTGTAATCATGAGTACTCTTGTAGACTGTACGAATATTTTGTACAGAGCAGCCTTTGGA 2255
```

```
QY 1681 AGGACTAGAGAAACTCTACAGGAGCCCAAGAGCTGCCTCCTGAGGAGAGAGCCAGAA 1740
Db 2256 AGGACTAGAGAAACTCTACAGGAGCCCAAGAGCTGCCTCCTGAGGAGAGAGCCAGAA 2315
QY 1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAA 1800
Db 2316 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAA 2375
QY 1801 CTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAAGACTCCAGGAACT 1860
Db 2376 CTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAAGACTCCAGGAACT 2435
QY 1861 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGATC 1920
Db 2436 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGATC 2495
QY 1921 CTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAACTCA 1980
Db 2496 CTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAACTCA 2555
QY 1981 GGCACCTCGAGGAGAAATTGGCGCTCTGAAAGAGAAAGCTGAGCCACGTCATGACCTTGC 2040
Db 2556 GGCACCTCGAGGAGAAATTGGCGCTCTGAAAGAGAAAGCTGAGCCACGTCATGACCTTGC 2615
QY 2041 TCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 2100
Db 2616 TCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 2675
QY 2101 C 2101
Db 2676 C 2676

RESULT 7
US-09-845-416-29
; Sequence 29, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 4825
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-29

Query Match 68.7%; Score, 1443; DB 13; Length 4825;
Best Local Similarity 84.6%; Pred. No. 0;
Matches 1777; Conservative 0; Mismatches 0; Indels 324; Gaps 1;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTACCACTCTGACCCCTACACGGAGCCCATT 60
Db 1657 GAGCTATGCCTACACACAGGCTGCTTATGTACCACTCTGACCCCTACACGGAGCCCATT 1716
QY 61 TCCTTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTTCATGATGAGAG 120
Db 1717 TCCTTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTTCATGATGAGAG 1776
QY 121 TGAAGTAAACCTGGACCGGTATCAACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC 180
Db 1777 TGAAGTAAACCTGGACCGGTATCAACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC 1836
QY 181 TGCTGAGGACACATTGCAAGCAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1837 TGCTGAGGACACATTGCAAGCAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAGA 1896
```

241 CCAGTTTCATACCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
1897 CCAGTTTCATACCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1956
301 TAAATATTTACAAATGGGAAGTAAGCTGATGGAAACAGGAAAATATATCAGAAGATGAAGA 360
1957 TAAATATTTACAAATGGGAAGTAAGCTGATGGAAACAGGAAAATATATCAGAAGATGAAGA 2016
361 AACTGAAGTACAAGAGCAGATGAATCTCCTAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
2017 AACTGAAGTACAAGAGCAGATGAATCTCCTAATTTCAAGATGGGAATGCCTCAGGGTAGC 2076
421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 480
2077 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 2136
481 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAACAGGAAAATGGAGGAAGA 540
2137 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAACAGGAAAATGGAGGAAGA 2196
541 GCCTCTTGGACCTGATCTTGAAGACCTAAACCGCCAAAGTACAACAACTAAGGTGCTTCA 600
2197 GCCTCTTGGACCTGATCTTGAAGACCTAAACCGCCAAAGTACAACAACTAAGGTGCTTCA 2256
601 AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATCTCTCACTCACATGGTGGTGTAGT 660
2257 AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATCTCTCACTCACATGGTGGTGTAGT 2316
661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTGGG 720
2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTGGG 2376
721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 780
2377 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGAC-- 2434
781 CCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 840
2435 ----- 2434
841 AAAAGAAGATGCAGTGAACAAAGATTCACAACTGGCTTTTAAAGATCAAAATGAAATGTT 900
2435 ----- 2434
901 ATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
2435 ----- 2434
961 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
2435 ----- 2434
1021 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1080
2435 ----- 2434
1081 ACTTGAAGAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 1140
2435 -----
1141 AAAGTTTCTTCCTGGCTTACAGAAGCTGAAACAACCTGCCAATGTCTTACAGGATGCTAC 1200
2473 AAAGTTTCTTCCTGGCTTACAGAAGCTGAAACAACCTGCCAATGTCTTACAGGATGCTAC 2532
1201 CCGTAAGGAAGAGGCTCCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 1260
2533 CCGTAAGGAAGAGGCTCCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 2592
1261 AGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAAACAG 1320
2593 AGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAAACAG 2652

1321 CCAAAAAATCCTGAGATCCCTGGAAAGTTCCGATGATGCAGTCTCTGTATCAAAAGACGTTT 1380
2653 CCAAAAAATCCTGAGATCCCTGGAAAGTTCCGATGATGCAGTCTCTGTATCAAAAGACGTTT 2712
1381 GGATAACATGAACCTCAAGTGGAGTGAATTCGGAAGAAAGTCTCTCAACATTAGGTCCCA 1440
2713 GGATAACATGAACCTCAAGTGGAGTGAATTCGGAAGAAAGTCTCTCAACATTAGGTCCCA 2772
1441 TTTGGAAGCCAGTTCTGACCAAGTGGAAAGCTGTGCACCTTTCTCTGAGGAACTTCTTGGT 1500
2773 TTTGGAAGCCAGTTCTGACCAAGTGGAAAGCTGTGCACCTTTCTCTGAGGAACTTCTTGGT 2832
1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGCGGACTTTCC 1560
2833 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGCGGACTTTCC 2892
1561 AGCAGTTTCAAGAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTAAGAACTAAAGA 1620
2893 AGCAGTTTCAAGAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTAAGAACTAAAGA 2952
1621 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGA 1680
2953 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGA 3012
1681 AGGACTAGAGAAACTCTPACCAAGGAGCCAGAGAGTGCCTCTCTGAGGAGAGAGCCAGAA 1740
3013 AGGACTAGAGAAACTCTPACCAAGGAGCCAGAGAGTGCCTCTCTGAGGAGAGAGCCAGAA 3072
1741 TGTCACTCGGCTTCTACGAAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAA 1800
3073 TGTCACTCGGCTTCTACGAAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAA 3132
1801 CCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 1860
3133 CCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 3192
1861 TCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATC 1920
3193 TCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATC 3252
1921 CTGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTTCAAGATCACCTCGAGAAAAGTCAA 1980
3253 CTGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTTCAAGATCACCTCGAGAAAAGTCAA 3312
1981 GGCACTTCGAGGAGAAATTCGCGCTCTGAAAAGAGAACGTGAGCCACGTCAATGACCTTGC 2040
3313 GGCACTTCGAGGAGAAATTCGCGCTCTGAAAAGAGAACGTGAGCCACGTCAATGACCTTGC 3372
2041 TCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTACCGTATAACCTCAGCACTCTGGAAGA 2100
3373 TCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTACCGTATAACCTCAGCACTCTGGAAGA 3432
2101 C 2101
3433 C 3433

RESULT 8

US-09-845-416-35
; Sequence 35, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35

| | | | | | |
|------------------|------|--|------|--|--|
| | | ; LENGTH: 4848 | | | |
| | | ; TYPE: DNA | | | |
| | | ; ORGANISM: Homo sapiens | | | |
| US-09-845-416-35 | | | | | |
| | | Query Match | | 68.7%; Score 1443; DB 13; Length 4848; | |
| | | Best Local Similarity | | 84.6%; Pred. No. 0; | |
| | | Matches 1777; Conservative | | 0; Mismatches 0; Indels 324; Gaps 1; | |
| QY | 1 | GAGCTATGCCTACACACAGGCTGCTTATGTACCACCTCTGACCCCTACACGGAGCCCAT | 60 | | |
| Db | 1680 | GAGCTATGCCTACACACAGGCTGCTTATGTACCACCTCTGACCCCTACACGGAGCCCAT | 1739 | | |
| QY | 61 | TCCTTACAGACATTTGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCATTGATGGAGAG | 120 | | |
| Db | 1740 | TCCTTACAGACATTTGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCATTGATGGAGAG | 1799 | | |
| QY | 121 | TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATTTATTCGTGGCTTCTTTTC | 180 | | |
| Db | 1800 | TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATTTATTCGTGGCTTCTTTTC | 1859 | | |
| QY | 181 | TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA | 240 | | |
| Db | 1860 | TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA | 1919 | | |
| QY | 241 | CCAGTTTCATCTCATGAGGGTACATGATGGAATTTGACAGCCCATCAGGGCCGGGTTGG | 300 | | |
| Db | 1920 | CCAGTTTCATCTCATGAGGGTACATGATGGAATTTGACAGCCCATCAGGGCCGGGTTGG | 1979 | | |
| QY | 301 | TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA | 360 | | |
| Db | 1980 | TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA | 2039 | | |
| QY | 361 | AAGTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC | 420 | | |
| Db | 2040 | AAGTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC | 2099 | | |
| QY | 421 | TAGCATGGAAAAACAAGCAATTTACATAGAGTTCCTAAGTCTCCAGATCAGAAACT | 480 | | |
| Db | 2100 | TAGCATGGAAAAACAAGCAATTTACATAGAGTTCCTAAGTCTCCAGATCAGAAACT | 2159 | | |
| QY | 481 | GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGGAAATGGAGGAAGA | 540 | | |
| Db | 2160 | GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGGAAATGGAGGAAGA | 2219 | | |
| QY | 541 | GCCTCTTGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACAATAAGGTGCTTCA | 600 | | |
| Db | 2220 | GCCTCTTGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACAATAAGGTGCTTCA | 2279 | | |
| QY | 601 | AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTTCTCACTCACATGGTGGTAGT | 660 | | |
| Db | 2280 | AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTTCTCACTCACATGGTGGTAGT | 2339 | | |
| QY | 661 | TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAAACAACCTTAAGGTATTGGG | 720 | | |
| Db | 2340 | TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAAACAACCTTAAGGTATTGGG | 2399 | | |
| QY | 721 | AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCCTTTTACAAGACAT | 780 | | |
| Db | 2400 | AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCCTTTTACAAGAC-- | 2457 | | |
| QY | 781 | CCTTCTCAAATGCCAACGCTCTTACTGAAGAACAGTGCCTTTTATGTGCAATGGCTTTTCA | 840 | | |
| Db | 2458 | ----- | 2457 | | |
| QY | 841 | AAAAGAAGATGCAGTGAACAAGATTCACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT | 900 | | |
| Db | 2458 | ----- | 2457 | | |
| QY | 901 | ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT | 960 | | |
| Db | 2458 | ----- | 2457 | | |

||||| 3396 TCGCCAGCTTACCACCTTTGGGCATTACAGCTCTCACCGTATTAACCTCAGCACTCTGGAAGA 3455
| 2101 C 2101
| 3456 C 3456

RESULT 9
JS-09-845-416-36
Sequence 36, Application US/09845416
Publication No. US20030171312A1
GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
TITLE OF INVENTION: THEREOF
FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 36
LENGTH: 5060
TYPE: DNA
ORGANISM: Homo sapiens
JS-09-845-416-36

Query Match 68.7%; Score 1443; DB 13; Length 5060;
Best Local Similarity 84.6%; Pred. No. 0;
Matches 1777; Conservative 0; Mismatches 0; Indels 324; Gaps 1;

2Y 1 GAGCTATGCCTACACACAGGCTGTATTATGTCACCACCTCTGACCCCTACACGGAGCCCAATT 60
| 1892 GAGCTATGCCTACACACAGGCTGTATTATGTCACCACCTCTGACCCCTACACGGAGCCCAATT 1951
| 61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
| 1952 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 2011
| 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTC 180
| 2012 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTC 2071
| 181 TGCTGAGGACACATTGCAAGCAACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
| 2072 TGCTGAGGACACATTGCAAGCAACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 2131
| 241 CCAGTTTCACTACTGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTGG 300
| 2132 CCAGTTTCACTACTGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTGG 2191
| 301 TAATATTCTACAATTGGGAAGTAAGCTGATTTGGAACAGGAAATATTATCAGAAGATGAAGA 360
| 2192 TAATATTCTACAATTGGGAAGTAAGCTGATTTGGAACAGGAAATATTATCAGAAGATGAAGA 2251
| 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
| 2252 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 2311
| 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGAATCTCCAGAATCAGAAACT 480
| 2312 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGAATCTCCAGAATCAGAAACT 2371
| 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAACAGGAAAAATGGAGGAAGA 540
| 2372 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAACAGGAAAAATGGAGGAAGA 2431
| 541 GCCTCTTGGACCTTGATCTTGAAGACCTTAAACGCCAAGTACAAACAACATAAGGTGCTTCA 600
| 2432 GCCTCTTGGACCTTGATCTTGAAGACCTTAAACGCCAAGTACAAACAACATAAGGTGCTTCA 2491

QY 601 AGAAGATCTAGAACAAAGCAAGTCAGGTCAAATTTCTCACTCACATGGTGGTGTAGT 660
| 2492 AGAAGATCTAGAACAAAGCAAGTCAGGTCAAATTTCTCACTCACATGGTGGTGTAGT 2551
| 661 TGATGAATCTAGTGGAGATCAGGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 720
| 2552 TGATGAATCTAGTGGAGATCAGGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 2611
| 721 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGTTCCTTTTACAAGACAT 780
| 2612 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGTTCCTTTTACAAGAC-- 2669
| 781 CCTTCTCAAATGGCAACCTCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGA 840
| 2670 --- 2669
| 841 AAAAGAAAGATGCAGTGAACAAGATTTCACAACTGGCTTTTAAAGATCAAAATGAAATGTT 900
| 2670 --- 2669
| 901 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAGCAATCCAT 960
| 2670 --- 2669
| 961 GGGCAAAACTGTATTCTACTCTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
| 2670 --- 2669
| 1021 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCCGCTGTGGGATAATTTAGTCCAAAA 1080
| 2670 --- 2669
| 1081 ACTTGAAGAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA 1140
| 2670 --- ACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA 2707
| 1141 AAAGTTTCTTGCCTGGCTTACAGAAAGCTGAAACAACCTGCAATGTCTACAGGATGCTAC 1200
| 2708 AAAGTTTCTTGCCTGGCTTACAGAAAGCTGAAACAACCTGCAATGTCTACAGGATGCTAC 2767
| 1201 CCGTAAGGAAAGGCTCCTAGAACACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCA 1260
| 2768 CCGTAAGGAAAGGCTCCTAGAACACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCA 2827
| 1261 AGACCTCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAAACAG 1320
| 2828 AGACCTCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAAACAG 2887
| 1321 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGAGTCCCTGTTACAAAGACGTTT 1380
| 2888 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGAGTCCCTGTTACAAAGACGTTT 2947
| 1381 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCA 1440
| 2948 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCA 3007
| 1441 TTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACTTTCTCTGCAGGAACCTTCTGCT 1500
| 3008 TTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACTTTCTCTGCAGGAACCTTCTGCT 3067
| 1501 GTGGCTACAGCTGAAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGGGCGACTTTCC 1560
| 3068 GTGGCTACAGCTGAAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGGGCGACTTTCC 3127
| 1561 AGCAGTTTCAGAAGCAGAACCGATGTACATAGGGCCTTCAAGGGGAATTGAAAACTAAAGA 1620
| 3128 AGCAGTTTCAGAAGCAGAACCGATGTACATAGGGCCTTCAAGGGGAATTGAAAACTAAAGA 3187
| 1621 ACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGA 1680
| 3188 ACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGA 3247
| 1681 AGGACTAGAGAAACTCTACCAGGAGGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCCAGAA 1740

| | | | |
|----|------|---|------|
| Db | 3248 | AGGACTAGAGAAACTCTACCGAGAGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAA | 3307 |
| QY | 1741 | TGTCACCTCGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAATTGAA | 1800 |
| Db | 3308 | TGTCACCTCGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAATTGAA | 3367 |
| QY | 1801 | CCTGCACCTCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCCAGGAACT | 1860 |
| Db | 3368 | CCTGCACCTCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCCAGGAACT | 3427 |
| QY | 1861 | TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATC | 1920 |
| Db | 3428 | TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATC | 3487 |
| QY | 1921 | CTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAA | 1980 |
| Db | 3488 | CTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAA | 3547 |
| QY | 1981 | GGCACTTCGAGAGAAAATTGGCGCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGC | 2040 |
| Db | 3548 | GGCACTTCGAGAGAAAATTGGCGCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGC | 3607 |
| QY | 2041 | TCGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATACCTCAGCACTCTGGAAGA | 2100 |
| Db | 3608 | TCGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATACCTCAGCACTCTGGAAGA | 3667 |
| QY | 2101 | C 2101 | |
| Db | 3668 | C 3668 | |

RESULT 10

US-10-149-736-42

Sequence 42, Application US/10149736

Publication No. US20030216332A1

GENERAL INFORMATION:

APPLICANT: Chamberlain, Jeffrey S.

APPLICANT: Harper, Scott Q.

TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences

FILE REFERENCE: UM-06968

CURRENT APPLICATION NUMBER: US/10/149,736

CURRENT FILING DATE: 2002-06-17

PRIOR APPLICATION NUMBER: PCT/US01/31126

PRIOR FILING DATE: 2001-10-04

PRIOR APPLICATION NUMBER: 60/238,848

PRIOR FILING DATE: 2000-10-06

NUMBER OF SEQ ID NOS: 96

SOFTWARE: PatentIn version 3.1

SEQ ID NO 42

LENGTH: 8689

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Synthetic

US-10-149-736-42

Query Match

Best Local Similarity

Matches 2101; Conservative

Score 1293; DB 13; Length 8689;

72.5%; Pred. No. 0;

0; Mismatches

0; Indels

798; Gaps

1;

| | | | |
|----|------|---|------|
| QY | 1 | GAGCTATGCCCTACACAGGCTGCTTATGTACACCACTCTGACCCCTACACGGAGCCCCATT | 60 |
| Db | 1099 | GAGCTATGCCCTACACAGGCTGCTTATGTACACCACTCTGACCCCTACACGGAGCCCCATT | 1158 |
| QY | 61 | TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG | 120 |
| Db | 1159 | TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG | 1218 |
| QY | 121 | TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC | 180 |
| Db | 1219 | TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC | 1278 |

| | | | |
|----|------|--|------|
| QY | 181 | TGCTGAGGACACATTGCAAGCACAAGGAGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA | 240 |
| Db | 1279 | TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA | 1338 |
| QY | 241 | CCAGTTTCATACATCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG | 300 |
| Db | 1339 | CCAGTTTCATACATCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG | 1398 |
| QY | 301 | TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTTATCAGAAGATGAAGA | 360 |
| Db | 1399 | TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTTATCAGAAGATGAAGA | 1458 |
| QY | 361 | AACTGAAGTACAAGAGCAGATGAATCTCTTAATTTCAAGATGGGAATGCCCTCAGGGTAGC | 420 |
| Db | 1459 | AACTGAAGTACAAGAGCAGATGAATCTCTTAATTTCAAGATGGGAATGCCCTCAGGGTAGC | 1518 |
| QY | 421 | TAGCATGGAAAAACAAAGCAATTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT | 480 |
| Db | 1519 | TAGCATGGAAAAACAAAGCAATTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT | 1578 |
| QY | 481 | GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGGAAAATGGAGGAAGA | 540 |
| Db | 1579 | GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGGAAAATGGAGGAAGA | 1638 |
| QY | 541 | GCCTCTTGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAAACATAGGTGCTTCA | 600 |
| Db | 1639 | GCCTCTTGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAAACATAGGTGCTTCA | 1698 |
| QY | 601 | AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTTCTCTACTCACATGGTGGTGGTAGT | 660 |
| Db | 1699 | AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTTCTCTACTCACATGGTGGTGGTAGT | 1758 |
| QY | 661 | TGATGAATCTAGTGGAGATCACGCCAACTGCTGCTTTGGAGAAACAACTTAAGGTATTGGG | 720 |
| Db | 1759 | TGATGAATCTAGTGGAGATCACGCCAACTGCTGCTTTGGAGAAACAACTTAAGGTATTGGG | 1818 |
| QY | 721 | AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGCTGGTCTTTTACAAGACAT | 780 |
| Db | 1819 | AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGCTGGTCTTTTACAAGACAT | 1878 |
| QY | 781 | CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTAGTGCAATGGCTTTCAGA | 840 |
| Db | 1879 | CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTAGTGCAATGGCTTTCAGA | 1938 |
| QY | 841 | AAAAGAAGATGCAGTGAAACAAGATTTCACAACTGGCTTTTAAAGATCAAAAATGAAATGTT | 900 |
| Db | 1939 | AAAAGAAGATGCAGTGAAACAAGATTTCACAACTGGCTTTTAAAGATCAAAAATGAAATGTT | 1998 |
| QY | 901 | ATCAAGTCTTCAAAAACTGGCCGCTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT | 960 |
| Db | 1999 | ATCAAGTCTTCAAAAACTGGCCGCTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT | 2058 |
| QY | 961 | GGGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGAC | 1020 |
| Db | 2059 | GGGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGAC | 2118 |
| QY | 1021 | CCAGAAGACGGAAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA | 1080 |
| Db | 2119 | CCAGAAGACGGAAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA | 2178 |
| QY | 1081 | ACTTGAAAAGAGTACAGCACAG | 1102 |
| Db | 2179 | ACTTGAAAAGAGTACAGCACAGATTTTACAGCAGCCCTGACCTAGCTCCTGGACTGACCAC | 2238 |
| QY | 1103 | ----- | 1102 |
| Db | 2239 | TATTGGAGCCTCTCCTACTCAGACTGTTACTCTGGTGACACAACTGTGGTTACTAAGGA | 2298 |
| QY | 1103 | ----- | 1102 |
| Db | 2299 | AACTGCCATCTCCAAAAC TAGAAATGCCATCTTCTTGATGTTGGAGGTACCTGCTCTGGC | 2358 |
| QY | 1103 | ----- | 1102 |

Db 2359 AGATTCAACCGGCTTGGACAGAACTTACCGACTGGCTTCTCTGCTGTATCAAGTTAT 2418
Qy 1103 ----- 1102
Db 2419 AAAATCACAGAGGGTGATGGTGGTGACCTTGAGGATATCAACGAGATGATCATCAAGCA 2478
Qy 1103 ----- 1102
Db 2479 GAAGGCAACAATGCAGGATTTGGAACAGAGGCGTCCCGAGTTGGAAGAACTCATTAACGC 2538
Qy 1103 ----- 1102
Db 2539 TGCCCAAAATTTGAAAAACAAGACCAGCAATCAAGAGGCTAGAACATCATTAACGGATCG 2598
Qy 1103 ----- 1102
Db 2599 AATTGAAAGAATTTCAGAAATCAGTGGGATGAAGTACAAGAACACCTTCAGAACCGGAGGCA 2658
Qy 1103 ----- 1102
Db 2659 ACAGTTGAATGAAATGTTAAAGGATTCAACACAAATGGCTGGAAGCTAAGGAAGAAGCTGA 2718
Qy 1103 ----- 1102
Db 2719 GCAGGTCTTAGGACAGGCCAGAGCCAAAGCTTGAGTCATGGAAGGAGGTCCCTATACAGT 2778
Qy 1103 ----- 1102
Db 2779 AGATGCAATCCAAAAAAGAAATCACAGAAACCAAGCAGTTGGCCAAAGACCTCCGCCAGTG 2838
Qy 1103 ----- 1102
Db 2839 GCAGACAAATGTAGATGTGGCAATGACTTTGGCCCTGAAACTTCTCCGGGATTATTCTGC 2898
Qy 1103 ----- 1102
Db 2899 AGATGATACCAGAAAAAGTCCACATGATAACAGAGAATATCAATGCCTCTTGGAGAAAGCAT 2958
Qy 1103 ----- ACTCATAGATTACTGCAACA 1122
Db 2959 TCATAAAAGGTTGAGTGAGCGAGAGGCTGCTTTGGAAGAACTCATAGATTACTGCAACA 3018
Qy 1123 GTTCCCTCCCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGAAAGTGAACAACACTGCCAA 1182
Db 3019 GTTCCCTCCCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGAAAGTGAACAACACTGCCAA 3078
Qy 1183 TGTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAGACTCCAAAGGGAGTAAAGA 1242
Db 3079 TGTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAGACTCCAAAGGGAGTAAAGA 3138
Qy 1243 GCTGATGAAACAATGGCAAGACCTCCAAAGTGAATTTGAAGCTCACACAGATGTTTTATCA 1302
Db 3139 GCTGATGAAACAATGGCAAGACCTCCAAAGTGAATTTGAAGCTCACACAGATGTTTTATCA 3198
Qy 1303 CAACCTGGATGAAAAACAGCCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGT 1362
Db 3199 CAACCTGGATGAAAAACAGCCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGT 3258
Qy 1363 CCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAGTC 1422
Db 3259 CCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAGTC 3318
Qy 1423 TCTCAACATTAGTTCCTCCATTTTGGAGCCAGTCTGACCAGTGGAAAGCGTCTGCACCTTTC 1482
Db 3319 TCTCAACATTAGTTCCTCCATTTTGGAGCCAGTCTGACCAGTGGAAAGCGTCTGCACCTTTC 3378
Qy 1483 TCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACC 1542
Db 3379 TCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACC 3438
Qy 1543 TATTGGAGGCGACTTTCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAG 1602

Db 3439 TATTGGAGGCGACTTTCCAGCAGTTCCAGAACAGAACGATGTACATAGGGCCTTCAAGAG 3498
Qy 1603 GGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTTGAGACTGTACGAATATTTCT 1662
Db 3499 GGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTTGAGACTGTACGAATATTTCT 3558
Qy 1663 GACAGAGCAGCCTTTGGAAGGACTAGAGAACTCTACCAAGGAGCCCAAGAGAGCTGCCTCC 1722
Db 3559 GACAGAGCAGCCTTTGGAAGGACTAGAGAACTCTACCAAGGAGCCCAAGAGAGCTGCCTCC 3618
Qy 1723 TGAGGAGAGAGCCCAAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATAC 1782
Db 3619 TGAGGAGAGAGCCCAAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATAC 3678
Qy 1783 TGAGTGGGAAAAATTTGAACCTTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCT 1842
Db 3679 TGAGTGGGAAAAATTTGAACCTTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCT 3738
Qy 1843 TGAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGC 1902
Db 3739 TGAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGC 3798
Qy 1903 TGAGGTGATCAAGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGA 1962
Db 3799 TGAGGTGATCAAGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGA 3858
Qy 1963 TCACCTCGAGAAAGTCAAGGCACCTCGAGGAGAAATTTGCGCTCTGAAAGAGAACGTGAG 2022
Db 3859 TCACCTCGAGAAAGTCAAGGCACCTCGAGGAGAAATTTGCGCTCTGAAAGAGAACGTGAG 3918
Qy 2023 CCACGTCAATGACCTTGTCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAA 2082
Db 3919 CCACGTCAATGACCTTGTCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAA 3978
Qy 2083 CCTCAGCACCTCTGGAAGAC 2101
Db 3979 CCTCAGCACCTCTGGAAGAC 3997

RESULT 11

US-09-845-416-32
; Sequence 32, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 4414
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-32

Query Match 53.0%; Score 1114; DB 13; Length 4414;
Best Local Similarity 97.0%; Pred. No. 2.3e-313;
Matches 1135; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCTACACGGAGCCCAAT 60
Db 1657 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCTACACGGAGCCCAAT 1716
Qy 61 TCCTTCACAGCATTTGGAAGCTCTGAAAGACAAGTCAATTTGGCAGTTCAATGATGGAGAG 120
Db 1717 TCCTTCACAGCATTTGGAAGCTCTGAAAGACAAGTCAATTTGGCAGTTCAATGATGGAGAG 1776
Qy 121 TGAAGTAAACCTGGACCGCTTATCAAAACAGCTTTTAGAAGAGTATTATCGTGGCTTCTTTC 180

Db 1777 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAAAGTATTATCGTGGCTTCTTTC 1836
QY 181 TGCTGAGGACACATTGCAAGCACAGGAGAGATTTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1837 TGCTGAGGACACATTGCAAGCACAGGAGAGATTTTCTAATGATGTGGAAGTGGTGAAGA 1896
QY 241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTCACAGCCCCATCAGGGCCGGGTTGG 300
Db 1897 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTCACAGCCCCATCAGGGCCGGGTTGG 1956
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1957 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 2016
QY 361 AACTCGAAGTACAAGAGCGAGATGAATCTCCTAAATTCGAAGATGGGAATGCCCTCAGGGTAGC 420
Db 2017 AACTGAAGTACAAGAGCGAGATGAATCTCCTAAATTCGAAGATGGGAATGCCCTCAGGGTAGC 2076
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 480
Db 2077 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 2136
QY 481 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAGAAAGAACCAAGGAAATGGAGGAAGA 540
Db 2137 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAGAAAGAACCAAGGAAATGGAGGAAGA 2196
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCAAAGTACAACAATGAAGGTGCTTCA 600
Db 2197 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCAAAGTACAACAATGAAGGTGCTTCA 2256
QY 601 AGAAGATCTAGAACCAAGAACAGTCAAGGTCAATTTCTCACTCACATGGTGGTGGTAGT 660
Db 2257 AGAAGATCTAGAACCAAGAACAGTCAAGGTCAATTTCTCACTCACATGGTGGTGGTAGT 2316
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACCAACTTAAGGTATTGGG 720
Db 2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACCAACTTAAGGTATTGGG 2376
QY 721 AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 780
Db 2377 AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 2436
QY 781 CCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTTTAGTGCATGGCTTTCAGA 840
Db 2437 CCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTTTAGTGCATGGCTTTCAGA 2496
QY 841 AAAAGAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT 900
Db 2497 AAAAGAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT 2556
QY 901 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 960
Db 2557 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 2616
QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTTCAACACTGAAGAAATAAGTCAGTGAC 1020
Db 2617 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTTCAACACTGAAGAAATAAGTCAGTGAC 2676
QY 1021 CCAGAAAGACGGAAGCATGGCTGGATAACTTTTGGCCGGTGTGGGATAAATTTAGTCCAAAA 1080
Db 2677 CCAGAAAGACGGAAGCATGGCTGGATAACTTTTGGCCGGTGTGGGATAAATTTAGTCCAAAA 2736
QY 1081 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCTCCCTGGACCTGGA 1140
Db 2737 ACTTGAAAAGAGTACAGCACAGACCCCTTGAAGAACTCCAGGAACCTTCAAGAGGCCACCGGA 2796
QY 1141 AAAGTTTCTTGCCTGGCTTACAGAAGCTGA 1170
Db 2797 TGAGCTGGACCTCAAGCTGCGCCCAAGCTGA 2826

RESULT 12

US-10-149-736-39
; Sequence 39, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 5417
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-39

Query Match 52.5%; Score 1103.4; DB 13; Length 5417;
Best Local Similarity 99.9%; Pred. No. 3.3e-310;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GAGCTATGCCTACACACAGCGTGCTTATGTACCACCTCTGTGACCTTACACGGAGCCCCATT 60
Db 1099 GAGCTATGCCTACACACAGCGTGCTTATGTACCACCTCTGTGACCTTACACGGAGCCCCATT 1158
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 120
Db 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGTTATCRAACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC 180
Db 1219 TGAAGTAAACCTGGACCGTTATCRAACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC 1278
QY 181 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1279 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
QY 241 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 300
Db 1339 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 1398
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 1458
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAAGAGCAGATGAATCTCCTTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 1518
QY 421 TAGCATGGAAAAACAAAGCAATTTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 480
Db 1519 TAGCATGGAAAAACAAAGCAATTTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 1578
QY 481 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAACCAAGGAAAAATGGAGGAAGA 540
Db 1579 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAACCAAGGAAAAATGGAGGAAGA 1638
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCAAAGTACAACAATAGGTGCTTCA 600
Db 1639 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCAAAGTACAACAATAGGTGCTTCA 1698
QY 601 AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTTCTCTCACTCACATGGTGGTGGTAGT 660
Db 1699 AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTTCTCTCACTCACATGGTGGTGGTAGT 1758
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACCACTTAAGGTATTGGG 720

Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAGAACAACCTTAAGGTATTGGG 1818
QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 780
Db 1819 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 1878
QY 781 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTATAGTCATGGCTTTCAGA 840
Db 1879 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTATAGTCATGGCTTTCAGA 1938
QY 841 AAAAGAAGATGCAGTGAACAAGATTCAACAACACTGGCTTTTAAAGATCAAAATGAAATGTT 900
Db 1939 AAAAGAAGATGCAGTGAACAAGATTCAACAACACTGGCTTTTAAAGATCAAAATGAAATGTT 1998
QY 901 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
Db 1999 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2058
QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGAC 2118
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAAATTTAGTCCAAAA 1080
Db 2119 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAAATTTAGTCCAAAA 2178
QY 1081 ACTTGAAAAGAGTACAGCACAGACT 1105
Db 2179 ACTTGAAAAGAGTACAGCACAGATT 2203

RESULT 13
US-10-149-736-44
; Sequence 44, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 11443
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic

US-10-149-736-44
Query Match 52.5%; Score 1103.4; DB 13; Length 11443;
Best Local Similarity 99.9%; Pred. No. 5.4e-310;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GAGCTATGCCCTACACACAGGCTGCTTATGTCACCACCTCTGACCTACACGGAGCCCAATT 60
Db 1099 GAGCTATGCCCTACACACAGGCTGCTTATGTCACCACCTCTGACCTACACGGAGCCCAATT 1158
QY 61 TCCTTCACAGCAATTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 120
Db 1159 TCCTTCACAGCAATTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCCTTC 180
Db 1219 TGAAGTAAACCTGGACCGGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCCTTC 1278

QY 181 TGCTGAGGACACATTGCAAGCACCAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 240
Db 1279 TGCTGAGGACACATTGCAAGCACCAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA*1338
QY 241 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTTGG 300
Db 1339 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTTGG 1398
QY 301 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1399 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 1458
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAAGAGCAGATGAATCTCTTAAATTCAGATGGGAATGCCTCAGGGTAGC 1518
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 480
Db 1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 1578
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAAACAAGGAAAAATGGAGGAAGA 540
Db 1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAAACAAGGAAAAATGGAGGAAGA 1638
QY 541 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAAACAACATAAGGTGCTTCA 600
Db 1639 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAAACAACATAAGGTGCTTCA 1698
QY 601 AGAAGATCTAGAAACAAGAACAGTCAGGGTCAATTTCTCACTCACATGGTGGTAGT 660
Db 1699 AGAAGATCTAGAAACAAGAACAGTCAGGGTCAATTTCTCACTCACATGGTGGTAGT 1758
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGSAGAACAACCTTAAGGTATTGGG 720
Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGSAGAACAACCTTAAGGTATTGGG 1818
QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGTGGGTTCTTTTACAAGACAT 780
Db 1819 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGTGGGTTCTTTTACAAGACAT 1878
QY 781 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTATAGTCATGGCTTTCAGA 840
Db 1879 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTATAGTCATGGCTTTCAGA 1938
QY 841 AAAAGAAGATGCAGTGAAACAAGATTCAACAACCTGGCTTTTAAAGATCAAAATGAAATGTT 900
Db 1939 AAAAGAAGATGCAGTGAAACAAGATTCAACAACCTGGCTTTTAAAGATCAAAATGAAATGTT 1998
QY 901 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
Db 1999 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2058
QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGAC 2118
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAAATTTAGTCCAAAA 1080
Db 2119 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAAATTTAGTCCAAAA 2178
QY 1081 ACTTGAAAAGAGTACAGCACAGACT 1105
Db 2179 ACTTGAAAAGAGTACAGCACAGATT 2203

RESULT 14
US-10-149-736-47
; Sequence 47, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences

FILE REFERENCE: UM-06968
CURRENT APPLICATION NUMBER: US/10/149,736
CURRENT FILING DATE: 2002-06-17
PRIORITY APPLICATION NUMBER: PCT/US01/31126
PRIORITY FILING DATE: 2001-10-04
PRIORITY APPLICATION NUMBER: 60/238,848
PRIORITY FILING DATE: 2000-10-06
NUMBER OF SEQ ID NOS: 96
SOFTWARE: PatentIn version 3.1
SEQ ID NO 47
LENGTH: 12057
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-10-149-736-47

Query Match 52.5%; Score 1103.4; DB 13; Length 12057;
Best Local Similarity 99.9%; Pred. No. 5.6e-310;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAATT 60
Db |||||
QY 1099 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAATT 1158
Db |||||
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGTATGGAGAG 120
Db |||||
QY 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGTATGGAGAG 1218
Db |||||
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGTATTATCGTGGCTTCTTTC 180
Db |||||
QY 1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGTATTATCGTGGCTTCTTTC 1278
Db |||||
QY 181 TGCTGAGCACACATTGCAAGCACAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 240
Db |||||
QY 1279 TGCTGAGCACACATTGCAAGCACAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 1338
Db |||||
QY 241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTTGG 300
Db |||||
QY 1339 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTTGG 1398
Db |||||
QY 301 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGAAAAATTATCAGAAATGAAGA 360
Db |||||
QY 1399 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGAAAAATTATCAGAAATGAAGA 1458
Db |||||
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
Db |||||
QY 1459 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518
Db |||||
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAAACT 480
Db |||||
QY 1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAAACT 1578
Db |||||
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAACAAAGGAAAAATGGAGGAAGA 540
Db |||||
QY 1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAACAAAGGAAAAATGGAGGAAGA 1638
Db |||||
QY 541 GCCTCTTGACCTGATCTTGAAGACCTAAAACGCCAAGTACAAACAACATTAAGGTGCTTCA 600
Db |||||
QY 1639 GCCTCTTGACCTGATCTTGAAGACCTAAAACGCCAAGTACAAACAACATTAAGGTGCTTCA 1698
Db |||||
QY 601 AGAAGATCTAGAACAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 660
Db |||||
QY 1699 AGAAGATCTAGAACAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 1758
Db |||||
QY 661 TGATGAATCTAGTGGAGATCACGCCACTGCTGCTTTTGGGAAGAACAACTTAAGGTATGGG 720
Db |||||
QY 1759 TGATGAATCTAGTGGAGATCACGCCACTGCTGCTTTTGGGAAGAACAACTTAAGGTATGGG 1818
Db |||||
QY 721 AGATCGATGGCAACATCTGTAGATGGACAGAACCCGCTGGGTTCTTTTACAGACAT 780
Db |||||
QY 1819 AGATCGATGGCAACATCTGTAGATGGACAGAACCCGCTGGGTTCTTTTACAGACAT 1878
Db |||||

QY 781 CCTTCTCAAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTATGTGCAATGGCTTTTCAGA 840
Db |||||
QY 1879 CCTTCTCAAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTATGTGCAATGGCTTTTCAGA 1938
Db |||||
QY 841 AAAAGAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTAAAGATCAAAAATGAAATGTT 900
Db |||||
QY 1939 AAAAGAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTAAAGATCAAAAATGAAATGTT 1998
Db |||||
QY 901 ATCAAGTCTTCAAAAACCTGGCGTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 960
Db |||||
QY 1999 ATCAAGTCTTCAAAAACCTGGCGTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 2058
Db |||||
QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db |||||
QY 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2118
Db |||||
QY 1021 CCAGAAGACGGAAGCATGGTGGATAACTTTGCCCGGTGTTGGGATAATTTTAGTCCAAAA 1080
Db |||||
QY 2119 CCAGAAGACGGAAGCATGGTGGATAACTTTGCCCGGTGTTGGGATAATTTTAGTCCAAAA 2178
Db |||||
QY 1081 ACTTGAAAAAGAGTACAGCACAGACT 1105
Db |||||
QY 2179 ACTTGAAAAAGAGTACAGCACAGACT 2203
Db |||||
RESULT 15
US-09-782-378A-22
; Sequence 22, Application US/09782378A
; Patent No. US20020102731A1
; GENERAL INFORMATION:
; APPLICANT: Hearing, Patrick
; APPLICANT: Bahou, Wadie
; APPLICANT: Sandalon, Ziv
; APPLICANT: Gnatenko, Dmitri
; TITLE OF INVENTION: Adenoviral Vectors
; FILE REFERENCE: STONYB-04970
; CURRENT APPLICATION NUMBER: US/09/782,378A
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/237,747
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 13957
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-378A-22

Query Match 52.5%; Score 1103.4; DB 10; Length 13957;
Best Local Similarity 99.9%; Pred. No. 6.2e-310;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAATT 60
Db |||||
QY 1099 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAATT 1158
Db |||||
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGTATGGAGAG 120
Db |||||
QY 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGTATGGAGAG 1218
Db |||||
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
Db |||||
QY 1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1278
Db |||||
QY 181 TGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGAAGA 240
Db |||||
QY 1279 TGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGAAGA 1338
Db |||||
QY 241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTTGG 300
Db |||||
QY 1339 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTTGG 1398
Db |||||
QY 301 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGAAAAATTATCAGAAATGAAGA 360
Db |||||

Db ||||| TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1458

2Y 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCCTCAGGTTAGC 420

Db |||||

1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCCTCAGGTTAGC 1518

QY 421 TAGCATGGAAAAACAAAGCAATTTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480

Db |||||

1519 TAGCATGGAAAAACAAAGCAATTTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 1578

QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAAACAAGGAAAAATGGAGGAAGA 540

Db |||||

1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAAACAAGGAAAAATGGAGGAAGA 1638

QY 541 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAAACAATAAGGTGCTTCA 600

Db |||||

1639 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAAACAATAAGGTGCTTCA 1698

2Y 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 660

Db |||||

1699 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 1758

2Y 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAGAACAACTTAAGGTATTGGG 720

Db |||||

1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAGAACAACTTAAGGTATTGGG 1818

2Y 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 780

Db |||||

1819 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 1878

2Y 781 CCTTCTCAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTAGTGCAATGGCTTTCAGA 840

Db |||||

1879 CCTTCTCAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTAGTGCAATGGCTTTCAGA 1938

2Y 841 AAAAGAGATGCAGTGAACAAGATTACACAACTGGCTTTAAAGATCAAAAATGAAATGTT 900

Db |||||

1939 AAAAGAGATGCAGTGAACAAGATTACACAACTGGCTTTAAAGATCAAAAATGAAATGTT 1998

2Y 901 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 960

Db |||||

1999 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 2058

2Y 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020

Db |||||

2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2118

2Y 1021 CCAGAGACGGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1080

Db |||||

2119 CCAGAGACGGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAA 2178

2Y 1081 ACTTGAAAAGAGTACAGCACAGACT 1105

Db |||||

2179 ACTTGAAAAGAGTACAGCACAGATT 2203

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 11:11:40 ; Search time 129.879 Seconds
(without alignments)
7140.092 Million cell updates/sec

Title: US-09-845-416-2_COPY_900_3000
Perfect score: 2101
Sequence: 1 gagctatgcctacacacagg.....acctcagcactctggaagac 2101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query % | | DB ID | Description |
|------------|--------|---------|---------|-------|---------------------|
| | | Match | Length | | |
| 1 | 1131 | 53.8 | 5952 | 4 | US-09-687-875A-1 |
| 2 | 1103.4 | 52.5 | 13977 | 4 | US-09-484-970B-60 |
| C 3 | 866.6 | 41.2 | 19307 | 3 | US-08-836-022A-10 |
| C 4 | 866.6 | 41.2 | 19307 | 3 | US-09-427-048A-10 |
| 5 | 343 | 16.3 | 6045 | 4 | US-09-091-501B-7 |
| 6 | 343 | 16.3 | 10320 | 4 | US-09-091-501B-9 |
| 7 | 79.4 | 3.8 | 200 | 4 | US-09-091-501B-5 |
| 8 | 78.6 | 3.7 | 200 | 4 | US-09-091-501B-4 |
| 9 | 78.6 | 3.7 | 200 | 4 | US-09-091-501B-6 |
| C 10 | 76.6 | 3.6 | 7218 | 1 | US-08-232-463-14 |
| 11 | 63.6 | 3.0 | 238 | 4 | US-09-687-875A-13 |
| C 12 | 44 | 2.1 | 1230025 | 4 | US-09-198-452A-1 |
| 13 | 43.4 | 2.1 | 1179 | 4 | US-09-107-532A-1186 |
| 14 | 40.4 | 1.9 | 2223 | 1 | US-08-257-073-4 |
| 15 | 39.2 | 1.9 | 16995 | 4 | US-08-961-527-82 |
| 16 | 38.6 | 1.8 | 1751 | 4 | US-09-620-312D-847 |
| 17 | 38.6 | 1.8 | 1995 | 1 | US-08-425-069-3 |
| 18 | 38.6 | 1.8 | 1995 | 2 | US-08-317-844B-3 |
| 19 | 38.2 | 1.8 | 1131 | 6 | 5180810-3 |
| 20 | 38.2 | 1.8 | 1784 | 6 | 5180810-2 |
| 21 | 38.2 | 1.8 | 1848 | 4 | US-09-134-001C-447 |
| 22 | 38.2 | 1.8 | 4929 | 4 | US-09-620-312D-674 |
| C 23 | 38 | 1.8 | 1394 | 4 | US-09-247-155-76 |
| C 24 | 37.4 | 1.8 | 2082 | 3 | US-08-985-335-4 |
| C 25 | 37.4 | 1.8 | 2082 | 3 | US-09-410-372-4 |
| C 26 | 37.2 | 1.8 | 2915 | 4 | US-09-336-115C-5 |
| C 27 | 37.2 | 1.8 | 3902 | 4 | US-08-961-527-212 |

| | | | | | | |
|----|------|-----|--------|---|--------------------|--------------------|
| 28 | 36.8 | 1.8 | 1886 | 6 | 5210183-1 | Patent No. 5210183 |
| 29 | 36.6 | 1.7 | 1845 | 4 | US-08-887-534A-22 | Sequence 22, Appl |
| 30 | 36.6 | 1.7 | 1845 | 4 | US-09-527-431-22 | Sequence 22, Appl |
| 31 | 36.2 | 1.7 | 289 | 3 | US-09-007-005-17 | Sequence 17, Appl |
| 32 | 36.2 | 1.7 | 289 | 3 | US-09-244-796-17 | Sequence 17, Appl |
| 33 | 36.2 | 1.7 | 2447 | 2 | US-09-014-969-14 | Sequence 14, Appl |
| 34 | 36.2 | 1.7 | 168575 | 4 | US-09-426-290-1 | Sequence 1, Appl |
| 35 | 36 | 1.7 | 608 | 3 | US-09-385-982-236 | Sequence 236, App |
| 36 | 36 | 1.7 | 2873 | 4 | US-08-630-915A-193 | Sequence 193, App |
| 37 | 35.8 | 1.7 | 790 | 3 | US-09-461-474-13 | Sequence 13, Appl |
| 38 | 35.6 | 1.7 | 4868 | 1 | US-08-139-937-12 | Sequence 12, Appl |
| 39 | 35.6 | 1.7 | 4868 | 5 | PCT-US93-11310-12 | Sequence 12, Appl |
| 40 | 35.6 | 1.7 | 7672 | 4 | US-09-220-132-24 | Sequence 24, Appl |
| 41 | 35.6 | 1.7 | 8257 | 4 | US-09-595-684B-30 | Sequence 30, Appl |
| 42 | 35.6 | 1.7 | 8789 | 1 | US-08-328-254-5 | Sequence 5, Appl |
| 43 | 35.6 | 1.7 | 10136 | 1 | US-08-353-700-2 | Sequence 2, Appl |
| 44 | 35.6 | 1.7 | 10136 | 5 | PCT-US95-16216-2 | Sequence 2, Appl |
| 45 | 35.2 | 1.7 | 3466 | 1 | US-08-468-036-38 | Sequence 38, Appl |

ALIGNMENTS

RESULT 1
US-09-687-875A-1
; Sequence 1, Application US/09687875A
; Patent No. 6544786
; GENERAL INFORMATION:
; APPLICANT: Xiao, Xiao
; APPLICANT: Liu, Paul
; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPLICED PE
; FILE REFERENCE: 00792
; CURRENT APPLICATION NUMBER: US/09/687,875A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/158,868
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 5952
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2897)..(2898)
; OTHER INFORMATION: S4 junction site
; NAME/KEY: misc feature
; LOCATION: (3198)..(3199)
; OTHER INFORMATION: S2 junction site
US-09-687-875A-1

Query Match 53.8%; Score 1131; DB 4; Length 5952;
Best Local Similarity 68.6%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 960; Gaps 1;

| | | | |
|----|------|---|------|
| Qy | 1 | GAGCTATGCCTACACAGGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCATTT | 60 |
| Db | 891 | GAGCTATGCCTACACAGGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCATTT | 950 |
| Qy | 61 | TCCTTCACAGCATTGGAGAGCTCCTGAAGACAGTCAATTTGGCAGTTCATTGATGGAGAG | 120 |
| Db | 951 | TCCTTCACAGCATTGGAGAGCTCCTGAAGACAGTCAATTTGGCAGTTCATTGATGGAGAG | 1010 |
| Qy | 121 | TGAAGTAAACCTGGACCGCTTATCAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTC | 180 |
| Db | 1011 | TGAAGTAAACCTGGACCGCTTATCAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTC | 1070 |
| Qy | 181 | TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAGTGGTGAAGA | 240 |
| Db | 1071 | TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAGTGGTGAAGA | 1130 |
| Qy | 241 | CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCGCGGTGG | 300 |

Db 1131 CCAGTTTCATACTCATGAGGGTACATGATGATTTGACAGCCCATCAGGCGGGTTGG 1190
Qy 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1191 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 1250
Qy 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
Db 1251 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1310
Qy 421 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 480
Db 1311 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 1370
Qy 481 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAAGAACAAAGGAAAAATGGAGGAAGA 540
Db 1371 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAAGAACAAAGGAAAAATGGAGGAAGA 1430
Qy 541 GCCTCTTTGGACCTGATCTTTGAAGACCTAAACGCCCAAGTACAAACAATAGGTGCTTCA 600
Db 1431 GCCTCTTTGGACCTGATCTTTGAAGACCTAAACGCCCAAGTACAAACAATAGGTGCTTCA 1490
Qy 601 AGAAGATCTAGAAACAAGAACRAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTAGT 660
Db 1491 AGAAGATCTAGAAACAAGAACRAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTAGT 1550
Qy 661 TGATGAATCTAGTGGAGATCAGCGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
Db 1551 TGATGAATCTAGTGGAGATCAGCGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1610
Qy 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGCCGCTGGTTCCTTTTACAAGACAT 780
Db 1611 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGCCGCTGGTTCCTTTTACAAGACAT 1670
Qy 781 CCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTACA 840
Db 1671 CCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTACA 1730
Qy 841 AAAAGAAGATGCAGTGAACAAAGATTACACAACACTGGCTTTTAAAGATCAAATGAATGTT 900
Db 1731 AAAAGAAGATGCAGTGAACAAAGATTACACAACACTGGCTTTTAAAGATCAAATGAATGTT 1790
Qy 901 ATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 960
Db 1791 ATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 1850
Qy 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db 1851 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1910
Qy 1021 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
Db 1911 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1970
Qy 1081 ACTTGAAAAAGAGTACAGCACAG----- 1102
Db 1971 ACTTGAAAAAGAGTACAGCACAGGAAACTGAAATAGCAGTTCAAGCTAAACACCGGATGT 2030
Qy 1103 ----- 1102
Db 2031 GGAAGAGATTTTGTCTAAAGGCGCAGCATTTGTACAGGAAAAACCAGCCACTCAGCCAGT 2090
Qy 1103 ----- 1102
Db 2091 GAAGAGGAAGTTAGAAGATCTGAGCTCTGAGTGGAGGCGGTAAACCGTTACTTCAAGA 2150
Qy 1103 ----- 1102
Db 2151 GCTGAGGGCAAGCAGCCTGACCTAGCTCCTGGACTGACCACCTATTGGAGCCTCTCCTAC 2210
Qy 1103 ----- 1102
Db 2211 TCAGACTGTTACTCTGGTGACACAACCTGTGTTTACTAAGGAAACTGCCATCTCCAAACT 2270

Qy 1103 ----- 1102
Db 2271 AGAAATGCCATCTTCTCTTGATGTTGGAGGTACCTGCTCTGGCAGATTTCAACCCGGGCTTG 2330
Qy 1103 ----- 1102
Db 2331 GACAGAACTTACCAGACTGGCTTTCTCTGTTGATCAAGTTATAAAATCACAGAGGGTGAT 2390
Qy 1103 ----- 1102
Db 2391 GGTGGTGACCTTGAGGATATCAACGAGATGATCATCAAGCAGAAGCAACAATGCAGGA 2450
Qy 1103 ----- 1102
Db 2451 TTTGGAAACAGAGGCGTCCCGAGTTGGAAGAACTCATTACCGCTGCCCAAATTTGAAAAA 2510
Qy 1103 ----- 1102
Db 2511 CAAGACCAGCAATCAAGAGGCTAGAAACAATCATTACGGATCGAATTGAAAGAAATTCAGAA 2570
Qy 1103 ----- 1102
Db 2571 TCAGTGGGATGAAGTACAAGAAACACCTTCAGAACCGGAGGCAACAGTTGAATGAAATGTT 2630
Qy 1103 ----- 1102
Db 2631 AAAGATTCAACACAATGGCTGGAAGCTAAGGAAGAAGCTGAGCAGGTCTTAGGACAGGC 2690
Qy 1103 ----- 1102
Db 2691 CAGAGCCAAGCTTGAGTCATGGAAGGAGGGTCCCTATACAGTAGATGCAATCCAAAAAGAA 2750
Qy 1103 ----- 1102
Db 2751 AATCACAGAAACCAAGCAGTTGGCCAAAGACCTCCGCCAGTGGCAGACAAAATGTAGATGT 2810
Qy 1103 ----- 1102
Db 2811 GGCAAAATGACTTGGCCCTGAAACTTCTCCGGGATTATTCTGCAGATGATACCAGAAAAAGT 2870
Qy 1103 ----- 1102
Db 2871 CCACATGATAACAGAGAATATCAATGCCTCTTGGAGAAGCATTTCTATAAAAGGTTGAGTGA 2930
Qy 1103 -----ACTCATAGATTACTGCAACAGTTTCCCTCCCTGGACCTGGA 1140
Db 2931 GCGAGAGGCTGCTTTTGGAGAAACTCATAGATTACTGCAACAGTTTCCCTCCCTGGACCTGGA 2990
Qy 1141 AAAATTTCTTGCTGCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTAC 1200
Db 2991 AAAATTTCTTGCTGCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTAC 3050
Qy 1201 CCGTAAGGAAAGGCTCTAGAAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 1260
Db 3051 CCGTAAGGAAAGGCTCTAGAAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 3110
Qy 1261 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAG 1320
Db 3111 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAG 3170
Qy 1321 CCAAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGACGTCTCTTACAAAGACGTTT 1380
Db 3171 CCAAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGACGTCTCTTACAAAGACGTTT 3230
Qy 1381 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCCA 1440
Db 3231 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCCA 3290
Qy 1441 TTTGGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTTCTCTGCAGGAACTTCTGGT 1500
Db 3291 TTTGGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTTCTCTGCAGGAACTTCTGGT 3350

```
2Y 1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCGGACTTTCC 1560
    |||||
Db 3351 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCGGACTTTCC 3410
    |||||
2Y 1561 AGCAGTTCAGAAAGCAGAACGATGTACATAGGGCTTCAAGAGGGAAATTGAAAACTAAAGA 1620
    |||||
Db 3411 AGCAGTTCAGAAAGCAGAACGATGTACATAGGGCTTCAAGAGGGAAATTGAAAACTAAAGA 3470
    |||||
2Y 1621 ACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGA 1680
    |||||
Db 3471 ACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGA 3530
    |||||
2Y 1681 AGGACTAGAGAAACTCTACACAGAGCCCGAGAGAGTGCCTCCTGAGGAGAGAGCCAGAA 1740
    |||||
Db 3531 AGGACTAGAGAAACTCTACACAGAGCCCGAGAGAGTGCCTCCTGAGGAGAGAGCCAGAA 3590
    |||||
2Y 1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAA 1800
    |||||
Db 3591 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAA 3650
    |||||
2Y 1801 CCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTTGAAAGACTCCAGGAAC 1860
    |||||
Db 3651 CCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTTGAAAGACTCCAGGAAC 3710
    |||||
2Y 1861 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATC 1920
    |||||
Db 3711 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATC 3770
    |||||
2Y 1921 CTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 1980
    |||||
Db 3771 CTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 3830
    |||||
2Y 1981 GGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGCTGAGCCACGTCATGACCTTGC 2040
    |||||
Db 3831 GGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGCTGAGCCACGTCATGACCTTGC 3890
    |||||
2Y 2041 TCGCCAGCTTACCACITTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 2100
    |||||
Db 3891 TCGCCAGCTTACCACITTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 3950
    |||||
2Y 2101 C 2101
    |
Db 3951 C 3951
```

RESULT 2

IS-09-484-970B-60
Sequence 60: Application US/09484970B
Patent No. 6426186

GENERAL INFORMATION:

APPLICANT: Jones, Karen A.
APPLICANT: Volkmuth, Wayne
APPLICANT: Walker, Michael G.
TITLE OF INVENTION: BONE REMODELING GENES
FILE REFERENCE: PB-0014 US
CURRENT APPLICATION NUMBER: US/09/484,970B
CURRENT FILING DATE: 2000-01-18
NUMBER OF SEQ ID NOS: 172
SOFTWARE: PERL Program
SEQ ID NO 60
LENGTH: 13977

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No. 6426186 229357.11CB1

NAME/KEY: unsure

LOCATION: 11721-11761, 12294, 13969

OTHER INFORMATION: a, t, c, g, or other

IS-09-484-970B-60

Query Match 52.5%; Score 1103.4; DB 4; Length 13977;
Best Local Similarity 99.9%; Pred. No. 0;

```
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GAGCTATGCCTACACACAGGCTGCTTATGTACCACCTCTGACCCCTACACGAGCCCCATT 60
    |||||
Db 1099 GAGCTATGCCTACACACAGGCTGCTTATGTACCACCTCTGACCCCTACACGAGCCCCATT 1158
    |||||
Qy 61 TCCTTTCACAGCATTGGAAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCAATGATGGAGAG 120
    |||||
Db 1159 TCCTTTCACAGCATTGGAAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCAATGATGGAGAG 1218
    |||||
Qy 121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTTGAAGAAAGTATTATCGTGGCTTCTTTC 180
    |||||
Db 1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTTGAAGAAAGTATTATCGTGGCTTCTTTC 1278
    |||||
Qy 181 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
    |||||
Db 1279 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
    |||||
Qy 241 CCAGTTTCATCTCATGAGGGGTACATGATGGAATTTGACAGCCCATCAGGGCCGGTTGG 300
    |||||
Db 1339 CCAGTTTCATCTCATGAGGGGTACATGATGGAATTTGACAGCCCATCAGGGCCGGTTGG 1398
    |||||
Qy 301 TAATATTCTACAATTGCGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 360
    |||||
Db 1399 TAATATTCTACAATTGCGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 1458
    |||||
Qy 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
    |||||
Db 1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 1518
    |||||
Qy 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 480
    |||||
Db 1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 1578
    |||||
Qy 481 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 540
    |||||
Db 1579 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAGGAAAAATGGAGGAAGA 1638
    |||||
Qy 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAATAAGGTGCTTCA 600
    |||||
Db 1639 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAATAAGGTGCTTCA 1698
    |||||
Qy 601 AGAAGATCTAGAACAAAGACAAGTCAGGTCAATTTCTCTCACTCAATGGTGGTAGT 660
    |||||
Db 1699 AGAAGATCTAGAACAAAGACAAGTCAGGTCAATTTCTCTCACTCAATGGTGGTAGT 1758
    |||||
Qy 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
    |||||
Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1818
    |||||
Qy 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGTCTTTTACAGACAT 780
    |||||
Db 1819 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGTCTTTTACAGACAT 1878
    |||||
Qy 781 CCTTCTCAAATGGCAACGTCCTTACTGAAGAACAGTGCCTTTTGTGTCATGGCTTTCAGA 840
    |||||
Db 1879 CCTTCTCAAATGGCAACGTCCTTACTGAAGAACAGTGCCTTTTGTGTCATGGCTTTCAGA 1938
    |||||
Qy 841 AAAAGAAGATGCAGTGAAACAAGATTCACACAACCTGGCTTTAAAGATCAAAAATGAAATGTT 900
    |||||
Db 1939 AAAAGAAGATGCAGTGAAACAAGATTCACACAACCTGGCTTTAAAGATCAAAAATGAAATGTT 1998
    |||||
Qy 901 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 960
    |||||
Db 1999 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 2058
    |||||
Qy 961 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
    |||||
Db 2059 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2118
    |||||
Qy 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1080
    |||||
Db 2119 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 2178
    |||||
```

QY 1081 ACTTGAAAAGAGTACAGCAGACT 1105
Db 2179 ACTTGAAAAGAGTACAGCAGATT 2203

RESULT 3

US-08-836-022A-10/c
; Sequence 10, Application US/08836022A
; Patent No. 6001557
; GENERAL INFORMATION:
; APPLICANT: Trustees of the University of Pennsylvania
; APPLICANT: Wilson, James M.
; APPLICANT: Fisher, Krishna J.
; APPLICANT: Chen, Shu-Jen
; APPLICANT: Weitzman, Matthew
; TITLE OF INVENTION: Improved Adenovirus Virus and
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P O Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,022A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/331,381
; FILING DATE: 28-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GNVPN.008PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19307 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
US-08-836-022A-10

Query Match 41.2%; Score 866.6; DB 3; Length 19307;
Best Local Similarity 86.5%; Pred. No. 6.1e-253;
Matches 956; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY 1 GAGCTATGCCTACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAT 60
Db 13568 GAGTTATGCCTTCACACAGGCTGCTTATGTTGCCACCTCTGATTCACACAGAGCCCTA 13509

QY 61 TCCTTCACAGCATTTGGAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCATTGATGGAGAG 120
Db 13508 TCCTTCACAGCATTTGGAGCTCCTGAGACAAAGTCAATTTGGCAGTTCATTGATGGAGAG 13449

QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
Db 13448 GGAAGTAAATCTGGATAGTTACCAAACTGCTTTAGAAGAAGTACTTTTCATGGCTTCTTTC 13389

QY 181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 13388 TGCCGAGGATACATTGCCGAGCACAAGGAGAGATTTCAAATGATGTTGAAGAAGTGAAGA 13329

QY 241 CCAGTTTTCATCTCATGAGGGGTACATCATGATGGATTGACAGCCCATCAGGGCCGGTTGG 300
Db 13328 ACAGTTTTCATCTCATGAGGGATTTCATGATGGATCTGACATCTCATCAAGGACTTGTGG 13269

QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTTGGAACAGGAAAATTATCAGAAGATGAAGA 360
Db 13268 TAATGTTCTACAGTTAGGAAGTCACTAGTTGGAAAAGGGAATATCAGAAGATGAAGA 13209

QY 361 AACTGAAGTACAGAGCAGATGAATCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
Db 13208 AGCTGAAGTGAAGAACAAATGAATCTCTCTAAATTTCAAGATGGGAATGTCTCAGGGTAGC 13149

QY 421 TAGCATGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480
Db 13148 TAGCATGAAAAACAAAGCAATTTACACAAAGTTCTAATGGATCTCCAGAATCAGAAACT 13089

QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGGAAAATGGAGGAAGA 540
Db 13088 AAAAGAACTAGATGACTGTTTAAACAAAACTGAAGAGAGAACTAAGAAAAATGGAGGAAGA 13029

QY 541 GCCTCTTGGACCTGATCTTTGAAGACCTTAAAAACGCCAAGTACAAACATAAGGTGCTTCA 600
Db 13028 GCCCTTTGGACCTGATCTTTGAAGATCTTAAATGCCAAGTACAAACATAAGGTGCTTCA 12969

QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTCCTCCTCAGTGGTGGTAGT 660
Db 12968 AGAAGATCTAGAACAGGAGCAGGTCAAGGTCACTCGTCACTCAGTGGTAGTGGT 12909

QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
Db 12908 TGATGAATCCAGCGGTGATCATGCAACAGCTGCTTTGGAAGAACAACTTAAGGTACTGGG 12849

QY 721 AGATCGATGGGCAAAACATCTGTAGTAGGACAGAACCGCTGGTCTTTTACAAGACAT 780
Db 12848 AGATCGATGGGCAAAATATCTGCAGATGGACTGAAGACCGCTGGATGTTTACAAGATAT 12789

QY 781 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTACA 840
Db 12788 TCTTCTAAATGGCAGCATTTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTACA 12729

QY 841 AAAAGAAGATGCAGTGAACAAAGATTCAACAACCTGGCTTTTAAAGATCAAAATGAAATGTT 900
Db 12728 AAAAGAAGATGCAATGAAGAACATTTCAGACAAAGTGGCTTTTAAAGATCAAAATGAAATGAT 12669

QY 901 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 960
Db 12668 GTCAAGTCTTCAAAAATATCTACTTTAAATAATAGATCTAGAAAAGAAAAAGCCCAACCAT 12609

QY 961 GGGCAAACTGTATTCACCTCAAAACAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db 12608 GGAATAAACTAAGTTCACTCAATCAAGATCTACTTTTCGGCAGTGAATAAATAAGTCAGTGAC 12549

QY 1021 CCAGAAGACGGGAAGCATGGCTGGATACCTTTGCCGGTGTGGGATAATTTAGTCCAAA 1080
Db 12548 TCAAAAGATGGAATCTGGATGGAATACTTTGCACAACGTTGGGACAATTTAACCACAAA 12489

QY 1081 ACTTGAAAAGAGTACAGCAGACT 1105
Db 12488 ACTTGAAAAGAGTTCAGCACAAATT 12464

RESULT 4

US-09-427-048A-10/c
; Sequence 10, Application US/09427048A
; Patent No. 6203975
; GENERAL INFORMATION:
; APPLICANT: Trustees of the University of Pennsylvania
; APPLICANT: Wilson, James M.
; APPLICANT: Fisher, Krishna J.
; APPLICANT: Chen, Shu-Jen
; APPLICANT: Weitzman, Matthew
; TITLE OF INVENTION: Improved Adenovirus Virus and
; Methods of Use Thereof

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, P O Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/427,048A
FILING DATE: 21-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/836,022
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNPVN.008PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
IS-09-427-048A-10
Query Match 41.2%; Score 866.6; DB 3; Length 19307;
Best Local Similarity 86.5%; Pred. No. 6.1e-253;
Matches 956; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

1 GAGCTATGCCTACACACAGGCTGCTTATGTACCACTCTGACCCCTACACGAGCCCAT 60
13568 GAGTTATGCCTTACACACAGGCTGCTTATGTTGCCACCTCTGATTCACACAGAGCCCCCTA 13509
61 TCCTTCACAGCATTTGGAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 120
13508 TCCTTCACAGCATTTGGAGCTCCACAGACAAGTCACCTTGACAGTTTCATTGATGGAGAG 13449
121 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAAGTATTATCTGGCTTCTTTC 180
13448 GGAAGTAAATCTGGATAGTTACCAACTGCTTTAGAGAAGTACTTTTCATGGCTTCTTTC 13389
181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAGTGGTGAAGA 240
13388 TGCCGAGGATACATTGGCAGCACAAGGAGAGATTTCAAATGATGTTGAAGAAGTGAAGA 13329
241 CCAGTTTCATCTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGCGCGGTTGG 300
13328 ACAGTTTCATGCTCATGAGGGATTTCATGATGGATCTGACATCTCATCAAGGACTTGTGG 13269
301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
13268 TAATGTTCTACAGTTAGGAAGTCAACTAGTTGGAAAAAGGGAATTTATCAGAAGATGAAGA 13209
361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCTCAGGGTAGC 420
13208 AGCTGAAGTGAAGAACAAATGAATCTCCTAAATTCAAGATGGGAATCTCTCAGGGTAGC 13149
421 TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480
13148 TAGCATGGAACAAACAAAGCAATTTACACAAAGTTCTAATGGATCTCCAGAATCAGAAATT 13089

QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGGAAAAATGGAGGAAGA 540
Db 13088 AAAAGAACTAGATGACTGGTTAAACAAAAAACTGAAGAGAGAACTAAGAAAAATGGAGGAAGA 13029
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCCAAGTACAACAACATAAGGTGCTTCA 600
Db 13028 GCCCTTGGACCTGATCTTGAAGATCTAAAAATGCCAAGTACAACAACATAAGGTGCTTCA 12969
QY 601 AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATCTCTCACTCACATGGTGGTGGTAGT 660
Db 12968 AGAAGATCTAGAACAGGAGCAGGTCAAGGTCAACTCGCTCACTCACATGGTGGTGGTAGT 12909
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATGGG 720
Db 12908 TGATGAATCCAGCGGTGATCATGCAACAGCTGCTTTGGAAGAACAACTTAAGGTATGGG 12849
QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGTTCCTTTTACAAGACAT 780
Db 12848 AGATCGATGGGCAAAATATCTGCAGATGGACTGAAGACCGCTGGATTGTTTACAAGATAT 12789
QY 781 CCTTCTCAAATGGCAACGCTTACTTGAAGAACAGTGCCTTTTGTGTCATGCTGCTTTCAGA 840
Db 12788 TCTTCTAAAATGGCAGCATTTTACTGAAGAACAGTGCCTTTTGTGTCATGCTGCTTTCAGA 12729
QY 841 AAAAGAAGATGCAGTGAACAAGATTCAACACAACCTGGCTTTAAAGATCAAAATGAAATGTT 900
Db 12728 AAAAGAAGATGCAATGAACAACATTTCAGACAAGTGGCTTTAAAGATCAAAATGAAATGAT 12669
QY 901 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 960
Db 12668 GTCAAGTCTTCAAAAAATATCTACTTTTAAANAATAGATCTAGAAAAAGAAAAAGCAACCAT 12609
QY 961 GGGCAAACTGTATTCTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db 12608 GGAANAACATAAGTTCACTCAATCAAGATCTACTTTGGCACTGAAAAATAAGTCAGTGAC 12549
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
Db 12548 TCAAAAAGATGGAATCTGATGGAAAACTTTTGCAACAACGTTGGGACAAATTAACCCAAAA 12489
QY 1081 ACTTGAANAAGAGTACAGCAGACT 1105
Db 12488 ACTTGAANAAGAGTTTCAGCACAAAT 12464

RESULT 5
US-09-091-501B-7
; Sequence 7, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Uterophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 6045
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(6037)

FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric
FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (724)..(758)
; OTHER INFORMATION: Precise residue is left open
US-09-091-501B-7

Query Match 16.3%; Score 343; DB 4; Length 6045;
Best Local Similarity 59.0%; Pred. No. 1.1e-93;
Matches 589; Conservative 0; Mismatches 410; Indels 0; Gaps 0;

Qy 104 AGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTA 163
Db 914 AGCACCGTCACTGAAGTGGACATGGATTGGACAGCTACCAGATAGCGCTAGAGGAAGTG 973

Qy 164 TTATCGTGGCTTCTTTCTGCTGAGGACACATTCGAAGCACAAGGAGAGATTCTTAATGAT 223
Db 974 CTGACGTGGCTGCTGTCGCGGAGGACACGTTCCAGAGCAAGATGACATTTCTGATGAT 1033

Qy 224 GTGGAAGTGGTGAAGACCACTTTCATCTACTCATGAGGGGTACATGATGGATTGACAGCC 283
Db 1034 GTCGAAGAAGTCAAGAGAGCAGTTTGCTACCCATGAACCTTTTATGATGGAGCTGACAGCA 1093

Qy 284 CATCAGGCCGGGTTGGTATATTTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAA 343
Db 1094 CACCAGAGCAGCGTGGGGAGCGCTCTGCAGGCTGGCAACCAGCTGATGACACAAGGGACT 1153

Qy 344 TTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGG 403
Db 1154 CTGTACAGAGGAGGAGGAGTTTGATGATCCAGGAACAGATGACCTTGCTGATGCAAGGTG 1213

Qy 404 GAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGAT 463
Db 1214 GAGCGCTCCGGGTGGAGAGCATGGAGAGGACAGTCCCGGCTGCACGACGCTCTGATGGAG 1273

Qy 464 CTCAGAATCAGAAACTGAAGAGATTGAATGACTGGCTAACAAAAACAGAAAGAAACA 523
Db 1274 CTGCAGAAGAAACAGCTGCAGCAGCTCTCAAGCTGGCTGGCCCTCACAGAAGAGCGCATT 1333

Qy 524 AGGAAATGGAGGAAGAGCCCTCTTGACCTGTATCTGAAGACCTTAAACGCCAAGTACAA 583
Db 1334 CAGAAGATGGAGAGCCCTCCGCTGGGTGATGACCTGCCCTCCCTGCAGAAGTGCTTCAA 1393

Qy 584 CAACATAAGGTGCTTCAAGAAGATCTAGAAGACATAAGAACAAAGTCAAGGTCAATCTCTCACT 643
Db 1394 GAACATAAAAAGTTTGCAAAATGACCTTGAAGCTGAACAGGTGAAGGTAAATTCCTTAACT 1453

Qy 644 CACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAA 703
Db 1454 CACATGGTGGTGAATTGTGGATGAAAAACAGTGGGGAGAGTGCCACAGCTCTCTGGAAGAT 1513

Qy 704 CAACCTAAGGTATTTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGG 763
Db 1514 CAGTTACAGAAACTGGGTGAGCGCTGGACAGCTGTATCCGCTGGACTGAAGAACGTTGG 1573

Qy 764 GTTCTTTTACAAGACATCCCTTCTCAAATGGCAACGTTCTACTGAAGAACAGTGCCTTTT 823
Db 1574 AACAGGTGCAAGAAATCAGTATCTGTGGCAGGAATATTGGAAGAGCAGTGTCTGTTG 1633

Qy 824 AGTGCATGGCTTTTCAGAAAAAGAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTAA 883
Db 1634 GAGGCTTGGCTCACCGAAAAAGGAAGAGGCTTTGATATAAGTTCAAACCCAGCAACTTTAA 1693

Qy 884 GATCAAAATGAAATGTTATCAAGTCTTCAAAAACCTGGCGTTTAAAAAGCGGATCTAGAA 943
Db 1694 GACCAGAAGGAACATAAGTGTCAAGTGTCCGGCGTCTGGCTATATTGAAGGAAGACATGGAA 1753

Qy 944 AAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCACACTG 1003
Db 1754 ATGAACAGGCAGACTCTGGATCAACTGAGTGAGATTGGCCAGGATGTGGGCCAATTACTC 1813

Qy 1004 AAGAATAAGTCAGTGACCCAGAGACGGAAGCATGGCTGGATAAATTTGGCCCGGTGTTGG 1063

Db 1814 AGTAATCCCAAGGCATCTAAGAAGATGAACAGTGAAGTCTCTGAGGAGCTAAACACAGAGATGG 1873

Qy 1064 GATAATTTAGTCCAAAAAAGTTGAAAAGAGTACACACAG 1102

Db 1874 GATTCTCTGGTTCAGAGACTCGAAGACTCTTCTTAACCAG 1912

RESULT 6
US-09-091-501B-9
; Sequence 9, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 10320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(10312)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Full length
; OTHER INFORMATION: utrophin construct
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (724)..(758)
; OTHER INFORMATION: Precise residue is left open
US-09-091-501B-9

Query Match 16.3%; Score 343; DB 4; Length 10320;
Best Local Similarity 59.0%; Pred. No. 1.1e-93;
Matches 589; Conservative 0; Mismatches 410; Indels 0; Gaps 0;

Qy 104 AGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTA 163
Db 914 AGCACCGTCACTGAAGTGGACATGGATTGGACAGCTACCAGATAGCGCTAGAGGAAGTG 973

Qy 164 TTATCGTGGCTTCTTTCTGCTGAGGACACATTCGAAGCACAAGGAGAGATTCTTAATGAT 223
Db 974 CTGACGTGGCTGCTGTCCGGGAGGACACCGTTCAGGAGCAAGATGACATTTCTGATGAT 1033

Qy 224 GTGGAAGTGGTGAAGACCACTTTCATCTACTCATGAGGGGTACATGATGGATTGACAGCC 283
Db 1034 GTCGAAGAAGTCAAGAGAGCAGTTTGCTACCCATGAACCTTTTATGATGGAGCTGACAGCA 1093

Qy 284 CATCAGGCCGGGTTGGTAAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAA 343
Db 1094 CACCAGAGCAGCGTGGGAGCGCTCTGCAGGCTGGCAACCAGCTGATGACACAAGGGACT 1153

Qy 344 TTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGG 403
Db 1154 CTGTACAGAGGAGGAGGAGTTTGATGATCCAGGAACAGATGACCTTGCTGATGCAAGGTG 1213

Qy 404 GAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGAT 463
Db 1214 GAGCGCTCCGGGTGGAGAGCATGGAGAGGACAGTCCCGGCTGCACGACGCTCTGATGGAG 1273

QY 464 CTCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAACA 523
DB 1274 CTGCAGAGAAACAGCTGCAGCAGCTCTCAAGCTGGCTGCAGAGAGCCGCAAT 1333
QY 524 AGGAAATGAGGAAGAGCCCTCTTGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAA 583
DB 1334 CAGAAGATGAGAGCCCTCCGCTGGGTGATGACCTGCCCTCCCTGCAGAAAGCTGCTCAA 1393
QY 584 CAACATAGGCTGCTTCAAGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTTCTCTCACT 643
DB 1394 GAACATATAAAGTTTGCAAAATGACCTTGAAGCTGAACAGGTGAAGTAAATTCCTTAAC 1453
QY 644 CACATGGTGGTGGTAGTTGATGAATCTAGTGGGCAACATCTGTAGATGGACAGAAGACCGCTGG 703
DB 1454 CACATGGTGGTGGTAGTTGGATGAACACAGTGGGAGAGTGCCACAGCTCTTCTGGAAGAT 1513
QY 704 CAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGCTGG 763
DB 1514 CAGTTACAGAAACTGGGTGAGCGCTGGACAGCTGTATGCCGCTGGACTGAAGAAGCTGG 1573
QY 764 GTTCTTTTACAAGACATCCTTCTCAAAATGGCAACGCTTACTGAAGAACAGTGGCTTTT 823
DB 1574 AACAGGTTGCAAGAAATCAGTATTCTGTGGCAGGAATTATTGGAAGAGCAGTGTCTGTG 1633
QY 824 AGTGCATGGCTTTCAGAAAAAGAGATGCAGTGAACACAGATTCACACAACTGGCTTTAAA 883
DB 1634 GAGGCTTGGCTCACCGAAAGGAAGAGGCTTTGAATAAGTTCAAAACCAGCAACTTTAA 1693
QY 884 GATCAAAATGAAATGTTATCAAGTCTTCAAAAACTGGCCGCTTTTAAAGCGGATCTAGAA 943
DB 1694 GACCAGAAGGAACCTAAGTGTCAAGTCCGGCTCTGGCTATATTGAAGGAAGACATGGAA 1753
QY 944 AAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAGATCTTCTTCAACACTG 1003
DB 1754 ATGAAGAGGCAGACTCTGGATCAACTGAGTGAGATTGSCCAGGATGTGGGCCAATTACTC 1813
QY 1004 AAGATAAAGTCAGTGACCCAGAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGG 1063
DB 1814 AGTAATCCCAAGGCATCTAAGAAGATGAACAGTGACTCTGAGGAGCTAACACAGAGATGG 1873
QY 1064 GATAATTTAGTCCAAAAACCTTGAAAAGAGTACAGCACAG 1102
DB 1874 GATTCTCTGGTTTCAGAGACTCGAAGACTCTTCTTAACCAG 1912

RESULT 7
US-09-091-501B-5
; Sequence 5, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Rattus sp.

US-09-091-501B-5
Query Match 3.8%; Score 79.4; DB 4; Length 200;
Best Local Similarity 62.4%; Pred. No. 4.2e-14;
Matches 123; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Best Local Similarity 64.3%; Pred. No. 2.4e-14;
Matches 119; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 565 CCTAAACGCCCAAGTACAACACATAAAGGTGCTTCAAGAGATCTAGAACAAAGAACAGT 624
DB 16 CCTGCAAAACCTGCTTGAAGAACATAAAAGTTTGCAAAGTGACCTCGAAGCTGAGCAGGT 75
QY 625 CAGGTCATTTCTCTCACTCACTCACTGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGC 684
DB 76 GAAGGTGAATTCCTTAACCTCATATGGTGGTGGTGGTGGATGAAGAACAGTGGGGAGAGCGC 135
QY 685 AACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAG 744
DB 136 CACAGCTGCTTTTGGAAAGATCAGTTACAGAAACTGGGTGAGCGCTGGACAGCTGTATGCCG 195
QY 745 ATGGA 749
DB 196 CTGGA 200

RESULT 8
US-09-091-501B-4
; Sequence 4, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Mus sp.

Query Match 3.7%; Score 78.6; DB 4; Length 200;
Best Local Similarity 62.4%; Pred. No. 4.2e-14;
Matches 123; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY 553 TGATCTTGAAGACCTAAACGCCCAAGTACAACACATAAAGGTGCTTCAAGAGATCTAGA 612
DB 4 TGACCTGCCCTCCCTGCAGAAAGCTGCTTCAAGAACATAAAAGTTTGCAAAATGACCTTGA 63
QY 613 ACAAGAACAAAGTCAGGTCATTTCTCTCACTCACTGCTGGTGGTGGTGGTGGTGGTGGTGG 672
DB 64 AGCTGAACAGGTGAAGGTAAATTCCTTAACCTCACTGCTGGTGGTGGTGGTGGTGGTGGTGG 123
QY 673 TGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGGAGATCGATGGGC 732
DB 124 TGGGAGAGTGCACAGCTCTTCTGGAAGATCAGTTACAGAACTGGGTGAGCGCTGGAC 183
QY 733 AACATCTGTAGATGGA 749
DB 184 AGCTGTATGCCGCTGGA 200

RESULT 9
US-09-091-501B-6
; Sequence 6, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:


```

RESULT 12
US-09-198-452A-1/c
; Sequence 1, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1
; LENGTH: 1230025
; TYPE: DNA
ORGANISM: Chlamydia pneumoniae
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(15000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (15001)..(30000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (30001)..(45000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (45001)..(60000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (60001)..(75000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (75001)..(90000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (90001)..(105000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (105001)..(120000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (120001)..(135000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (135001)..(150000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (150001)..(165000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (165001)..(180000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (180001)..(195000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature

```


[illegible]

RESULT 14

JS-08-257-073-4
: Sequence 4, Application US/08257073
: Patent No. 5766597
: GENERAL INFORMATION:
: APPLICANT: Paoletti, Enzo
: APPLICANT: de Taisne, Charles
: APPLICANT: Tine, John A.
: TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE
: NUMBER OF SEQUENCES: 143
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Curtis, Morris & Safford, P.C.
: STREET: 530 Fifth Avenue, 25th Floor
: CITY: New York
: STATE: New York
: COUNTRY: UNITED STATES OF AMERICA
: ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US 08/257,073
 FILING DATE: 09-JUN-1994
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/075,783
 FILING DATE: 11-JUN-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/852,305
 FILING DATE: 18-MAR-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/672,183
 FILING DATE: 20-MAR-1991

ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,5
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
TELEX: 425066 CURTMS
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2223 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

[illegible]

RESULT 15

US-08-961-527-82
; Sequence 82, Application US/08961527

Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

```
;
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1
;
; COMPUTER: HP Vectra 486/33
;
; OPERATING SYSTEM: MSDOS version 6.2
;
; SOFTWARE: ASCII Text
;
```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

```

Search completed: February 2, 2004, 06:15:45
Job time : 145.879 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 10:13:25 ; Search time 499.045 Seconds
(without alignments)
11364.749 Million cell updates/sec

Title: US-09-845-416-2_COPY_900_3000
Perfect score: 2101
Sequence: 1 gagctatgcctacacacagg.....acctcagcaactctggaagac 2101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----------|--------------------|
| 1 | 2101 | 100.0 | 4182 | AAD37230 | Human dystrophin m |
| 2 | 2101 | 100.0 | 5149 | AAD37255 | Adeno-associated v |
| 3 | 1579.8 | 75.2 | 3999 | AAD37234 | Human dystrophin m |
| 4 | 1579.8 | 75.2 | 4966 | AAD37256 | Adeno-associated v |
| 5 | 1579.8 | 75.2 | 4990 | AAD37262 | Adeno-associated v |
| 6 | 1443 | 68.7 | 3858 | AAD37237 | Human dystrophin m |
| 7 | 1443 | 68.7 | 4825 | AAD37257 | Adeno-associated v |
| 8 | 1443 | 68.7 | 4848 | AAD37263 | Adeno-associated v |

| | | | | | | |
|----|--------|------|-------|----|-----------|--------------------|
| 9 | 1443 | 68.7 | 5060 | 24 | AAD37264 | Adeno-associated v |
| 10 | 1293 | 61.5 | 8689 | 24 | ABK82000 | DNA encoding mini- |
| 11 | 1131 | 53.8 | 5952 | 22 | AAD06794 | Human dystrophin g |
| 12 | 1114 | 53.0 | 4414 | 24 | AAD37260 | Adeno-associated v |
| 13 | 1103.4 | 52.5 | 5417 | 24 | ABK81997 | DNA encoding mini- |
| 14 | 1103.4 | 52.5 | 11241 | 24 | ABK82005 | CDNA encoding huma |
| 15 | 1103.4 | 52.5 | 11443 | 24 | ABK82002 | DNA encoding mini- |
| 16 | 1103.4 | 52.5 | 12923 | 10 | AAAN90338 | Sequence of human |
| 17 | 1103.4 | 52.5 | 13957 | 24 | ABT10904 | Human breast cance |
| 18 | 1103.4 | 52.5 | 13957 | 24 | ABS69900 | Human dystrophin g |
| 19 | 1103.4 | 52.5 | 13957 | 24 | ABN95786 | Gene #2284 used to |
| 20 | 1103.4 | 52.5 | 13957 | 24 | ABK81959 | CDNA encoding huma |
| 21 | 1103.4 | 52.5 | 13977 | 24 | ABS70403 | Human bone remodel |
| 22 | 1102 | 52.5 | 3446 | 24 | AAD37242 | Human dystrophin m |
| 23 | 1100.2 | 52.4 | 11058 | 24 | AAD37229 | Human dystrophin p |
| 24 | 1097.8 | 52.3 | 1991 | 24 | AAD37231 | Human dystrophin N |
| 25 | 999 | 47.5 | 2169 | 24 | AAD37232 | Human dystrophin r |
| 26 | 999 | 47.5 | 3531 | 24 | AAD37238 | Human dystrophin m |
| 27 | 999 | 47.5 | 4498 | 24 | AAD37258 | Adeno-associated v |
| 28 | 997 | 47.5 | 5339 | 24 | ABK81998 | DNA encoding mini- |
| 29 | 996 | 47.4 | 5462 | 24 | ABK81999 | DNA encoding mini- |
| 30 | 911.6 | 43.4 | 4402 | 21 | AAZ48568 | A rod shortened dy |
| 31 | 866.6 | 41.2 | 13815 | 24 | ABK81960 | CDNA encoding mous |
| 32 | 866.6 | 41.2 | 13815 | 24 | ABI99799 | Mouse ischaemic co |
| 33 | 866.6 | 41.2 | 19307 | 17 | AAT27558 | Shuttle vector pAD |
| 34 | 858.6 | 40.9 | 13815 | 19 | AAV18885 | Mus musculus dyst |
| 35 | 855.4 | 40.7 | 3275 | 10 | AAAN97129 | Partial sequence o |
| 36 | 787 | 37.5 | 3510 | 24 | AAD37240 | Human dystrophin m |
| 37 | 787 | 37.5 | 4476 | 24 | AAD37259 | Adeno-associated v |
| 38 | 777 | 37.0 | 1667 | 24 | AAD37235 | Human dystrophin N |
| 39 | 718.8 | 34.2 | 4075 | 21 | AAZ48569 | A rod shortened dy |
| 40 | 714.6 | 34.0 | 4402 | 21 | AAZ48567 | A rod shortened dy |
| 41 | 652 | 31.0 | 1821 | 24 | AAD37241 | Human dystrophin r |
| 42 | 450 | 21.4 | 1340 | 24 | AAD37239 | Human dystrophin N |
| 43 | 409.2 | 19.5 | 3747 | 21 | AAZ48566 | A rod shortened dy |
| 44 | 387 | 18.4 | 387 | 24 | ABK81990 | Human dystrophin s |
| 45 | 361.4 | 17.2 | 10302 | 24 | ABS69901 | Human utrophin gen |

ALIGNMENTS

RESULT 1

| | |
|----------|---|
| AAD37230 | |
| ID | AAD37230 standard; DNA; 4182 BP. |
| XX | |
| AC | AAD37230; |
| XX | |
| DT | 21-AUG-2002 (first entry) |
| XX | |
| DE | Human dystrophin minigene delta4173. |
| XX | |
| KW | Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin; |
| KW | adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD; |
| KW | Becker muscular dystrophy; ds. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| PN | WO200183695-A2. |
| XX | |
| PD | 08-NOV-2001. |
| XX | |
| PF | 27-APR-2001; 2001WO-US13677. |
| XX | |
| PR | 28-APR-2000; 2000US-200777P. |
| XX | |
| PA | (XIAO/) XIAO X. |
| XX | |
| PI | Xiao X; |
| XX | |
| DR | WPI; 2002-049342/06. |
| XX | |
| PT | New dystrophin minigene for treating Duchenne or Becker muscular |

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 43-44; 71pp; English.
XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta4173 containing nucleotides 1-1992 (N-terminus,
CC hinge H1 and rods R1, R2 and R3), 8059-10227 (rods R22, R23 and R24,
CC hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids).
XX
SQ Sequence 4182 BP; 1309 A; 927 C; 970 G; 976 T; 0 other;
Query Match 100.0%; Score 2101; DB 24; Length 4182;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAT 60
DB 900 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAT 959
QY 61 TCCTTCACAGCATTTGGAGCTCCTGAAGACAGATTTTGGCAGTTTCAATGATGGAGAG 120
DB 960 TCCTTCACAGCATTTGGAGCTCCTGAAGACAGATTTTGGCAGTTTCAATGATGGAGAG 1019
QY 121 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
DB 1020 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1079
QY 181 TGCTGAGGACACATTGGGAAGTAAGCTGATTGGAAACAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
DB 1080 TGCTGAGGACACATTGGCAAGCACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1139
QY 241 CCAGTTTCATCTCATGAGGCTACATGATGATTTGACAGCCCATCAGGCGCGGTTGG 300
DB 1140 CCAGTTTCATCTCATGAGGCTACATGATGATTTGACAGCCCATCAGGCGCGGTTGG 1199
QY 301 TAATATCTCAATTTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 360
DB 1200 TAATATCTCAATTTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 1259
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTTAATTCAGATGGGAATGCCTCAGGGTAGC 420
DB 1260 AACTGAAGTACAAGAGCAGATGAATCTCTTAATTCAGATGGGAATGCCTCAGGGTAGC 1319
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 480
DB 1320 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 1379
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGGAAATGGAGGAAGA 540
DB 1380 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGGAAATGGAGGAAGA 1439
QY 541 GCCTCTTGACCTGATCTTGAAGACCTAAAAACGCAAGTACAAACATAGGTGCTTCA 600
DB 1440 GCCTCTTGACCTGATCTTGAAGACCTAAAAACGCAAGTACAAACATAGGTGCTTCA 1499
QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACTCATGGTGGTAGT 660
DB 1500 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACTCATGGTGGTAGT 1559
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGGAAGAACAACTTAAGGTATTTGGG 720
DB 1560 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGGAAGAACAACTTAAGGTATTTGGG 1619

QY 721 AGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGTCTTTTACAAGACAT 780
DB 1620 AGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGTCTTTTACAAGACAT 1679
QY 781 CCTTCTCAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTTTACTGTCATGGCTTTTACA 840
DB 1680 CCTTCTCAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTTTACTGTCATGGCTTTTACA 1739
QY 841 AAAAGAAGATGAGTGAACAAAGATTTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 900
DB 1740 AAAAGAAGATGAGTGAACAAAGATTTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 1799
QY 901 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 960
DB 1800 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 1859
QY 961 GGGCAAACTGTATCTACTCAAAACAAGATCTTTTCAACACACTGAAGATAAGTCAGTGAC 1020
DB 1860 GGGCAAACTGTATCTACTCAAAACAAGATCTTTTCAACACACTGAAGATAAGTCAGTGAC 1919
QY 1021 CCAGAAGACCGGAAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTTAGTCCAAAA 1080
DB 1920 CCAGAAGACCGGAAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTTAGTCCAAAA 1979
QY 1081 ACTTGAAGAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA 1140
DB 1980 ACTTGAAGAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA 2039
QY 1141 AAAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTACAGGATGCTAC 1200
DB 2040 AAAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTACAGGATGCTAC 2099
QY 1201 CCGTAAGGAAAGGCTCTAGAAAGACTCCAGAGAGTAAAGAGCTGATGAAACAATGGCA 1260
DB 2100 CCGTAAGGAAAGGCTCTAGAAAGACTCCAGAGAGTAAAGAGCTGATGAAACAATGGCA 2159
QY 1261 AGACCTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCAACACCTGGATGAAACAG 1320
DB 2160 AGACCTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCAACACCTGGATGAAACAG 2219
QY 1321 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGTCAGTCTCTGAGGAGCGTCTT 1380
DB 2220 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGTCAGTCTCTGAGGAGCGTCTT 2279
QY 1381 GGAATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCCA 1440
DB 2280 GGAATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCCA 2339
QY 1441 TTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGACACCTTTCTCTGAGGAACTTCTGGT 1500
DB 2340 TTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGACACCTTTCTCTGAGGAACTTCTGGT 2399
QY 1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTTCC 1560
DB 2400 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTTCC 2459
QY 1561 AGCAGTTCAAGAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTTGAAAACTAAAAA 1620
DB 2460 AGCAGTTCAAGAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTTGAAAACTAAAAA 2519
QY 1621 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGA 1680
DB 2520 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGA 2579
QY 1681 AGGACTAGAGAAAACTTACAGGAGCCCGAGAGCTGCCTCCTGAGGAGAGAGCCAGAA 1740
DB 2580 AGGACTAGAGAAAACTTACAGGAGCCCGAGAGCTGCCTCCTGAGGAGAGAGCCAGAA 2639
QY 1741 TGTCACTGGGCTTCTACGAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAA 1800
DB 2640 TGTCACTGGGCTTCTACGAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAA 2699

QY 1801 CCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 1860
|||||
Db 2700 CCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 2759
QY 1861 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGGCCAAGCTGAGGTGATCAAGGGATC 1920
|||||
Db 2760 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGGCCAAGCTGAGGTGATCAAGGGATC 2819
QY 1921 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 1980
|||||
Db 2820 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 2879
QY 1981 GGCACCTTCGAGGAGAAATTGCGCCTCTGAAAAGAGAACGTGAGCCACGTCAATGACCTTGC 2040
|||||
Db 2880 GGCACCTTCGAGGAGAAATTGCGCCTCTGAAAAGAGAACGTGAGCCACGTCAATGACCTTGC 2939
QY 2041 TCGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 2100
|||||
Db 2940 TCGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 2999
QY 2101 C 2101
Db 3000 C 3000

RESULT 2
AAD37255
ID AAD37255 standard; DNA; 5149 BP.
AC AAD37255;
XX
JT 21-AUG-2002 (first entry)
DE Adeno-associated virus vector plasmid, AAV-MCK-delta4173.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
DS Chimeric - Homo sapiens.
DS Chimeric - Unidentified.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 57-59; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV

CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence.
XX
SQ Sequence 5149 BP; 1489 A; 1236 C; 1269 G; 1155 T; 0 other;
Query Match 100.0%; Score 2101; DB 24; Length 5149;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACAGCTTTAGAGAAGTATTATCGTGGCTTCTTC 60
Db 1657 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCTACACGGAGCCCAT 1716
QY 61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 120
Db 1717 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 1776
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTC 180
Db 1777 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTC 1836
QY 181 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1837 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1896
QY 241 CCAGTTTCATACATGAGGGGTACATGATGGATTGACAGCCCCATCAGGGCCGGGTGG 300
Db 1897 CCAGTTTCATACATGAGGGGTACATGATGGATTGACAGCCCCATCAGGGCCGGGTGG 1956
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1957 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 2016
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGCTAGC 420
Db 2017 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGCTAGC 2076
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 480
Db 2077 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 2136
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAACAAGGAAAAATGGAGGAAGA 540
Db 2137 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAACAAGGAAAAATGGAGGAAGA 2196
QY 541 GCCTCTTGGACCTGATCTTTGAAGACCTAAACGCCAACTACAAACAATAAGGTGCTTCA 600
Db 2197 GCCTCTTGGACCTGATCTTTGAAGACCTAAACGCCAACTACAAACAATAAGGTGCTTCA 2256
QY 601 AGAAGATCTAGAACCAAGAACAGTCAAGGTCAATTCTCTCACTCACATGGTGGTGTAGT 660
Db 2257 AGAAGATCTAGAACCAAGAACAGTCAAGGTCAATTCTCTCACTCACATGGTGGTGTAGT 2316
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
Db 2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 2376
QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTCTTTTACAAACAT 780
Db 2377 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTCTTTTACAAACAT 2436
QY 781 CTTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 840
Db 2437 CTTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 2496
QY 841 AAAAGAAGATGCAGTGAACAAGATTCAACAACCTGGCTTTAAAGATCAAAATGAATGTT 900
Db 2497 AAAAGAAGATGCAGTGAACAAGATTCAACAACCTGGCTTTAAAGATCAAAATGAATGTT 2556
QY 901 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 960
Db 2557 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 2616

QY 961 GGGCAAACGTGATTCACCTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db |||||
QY 2617 GGGCAAACGTGATTCACCTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2676
Db |||||
QY 1021 CCAGAAGACGGAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTAGTCCAAAA 1080
Db |||||
QY 2677 CCAGAAGACGGAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTAGTCCAAAA 2736
Db |||||
QY 1081 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA 1140
Db |||||
QY 2737 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA 2796
Db |||||
QY 1141 AAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACACTGCCAATGTCTCTACAGGATGCTAC 1200
Db |||||
QY 2797 AAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACACTGCCAATGTCTCTACAGGATGCTAC 2856
Db |||||
QY 1201 CCGTAAGGAAAGGCTCCTAGAAAGACTCCAAAGGGAGTAAAGAGCTGATGAAACAATGGCA 1260
Db |||||
QY 2857 CCGTAAGGAAAGGCTCCTAGAAAGACTCCAAAGGGAGTAAAGAGCTGATGAAACAATGGCA 2916
Db |||||
QY 1261 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAAACAG 1320
Db |||||
QY 2917 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAAACAG 2976
Db |||||
QY 1321 CCAAAAAATCCTGAGATCCCTGGAGGTTCCGATGATGCAGTCCCTGTTACAAAGACGTTT 1380
Db |||||
QY 2977 CCAAAAAATCCTGAGATCCCTGGAGGTTCCGATGATGCAGTCCCTGTTACAAAGACGTTT 3036
Db |||||
QY 1381 GGATAACATGAACCTCAAGTGGAGTGAACTTCGGAATAAGTCTCTCAACATTAGGTCCCA 1440
Db |||||
QY 3037 GGATAACATGAACCTCAAGTGGAGTGAACTTCGGAATAAGTCTCTCAACATTAGGTCCCA 3096
Db |||||
QY 1441 TTTGGAAGCCAGTTCTGACCAAGTGGAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGT 1500
Db |||||
QY 3097 TTTGGAAGCCAGTTCTGACCAAGTGGAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGT 3156
Db |||||
QY 1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCGCACTTTCC 1560
Db |||||
QY 3157 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCGCACTTTCC 3216
Db |||||
QY 1561 AGCAGTTCAGAAGCAGAACCATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 1620
Db |||||
QY 3217 AGCAGTTCAGAAGCAGAACCATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 3276
Db |||||
QY 1621 ACCTGTAATCATGATGATCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGA 1680
Db |||||
QY 3277 ACCTGTAATCATGATGATCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGA 3336
Db |||||
QY 1681 AGGACTAGAGAACTCTACAGGAGCCAGAGAGTGCTCCTCTGAGGAGAGAGCCAGAA 1740
Db |||||
QY 3337 AGGACTAGAGAACTCTACAGGAGCCAGAGAGTGCTCCTCTGAGGAGAGAGCCAGAA 3396
Db |||||
QY 1741 TGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAA 1800
Db |||||
QY 3397 TGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAA 3456
Db |||||
QY 1801 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAAC 1860
Db |||||
QY 3457 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAAC 3516
Db |||||
QY 1861 TCAAGAGCCACGGATGAGCTGGACCTCAAGCTCGGCCAAGCTGAGGTGATCAAGGGATC 1920
Db |||||
QY 3517 TCAAGAGCCACGGATGAGCTGGACCTCAAGCTCGGCCAAGCTGAGGTGATCAAGGGATC 3576
Db |||||
QY 1921 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 1980
Db |||||
QY 3577 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 3636
Db |||||
QY 1981 GGCACCTTCAGAGGAGAAATTGGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 2040
Db |||||
QY 3637 GGCACCTTCAGAGGAGAAATTGGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 3696
Db |||||
QY 2041 TCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCAACCGTATACCTCAGCACCTCTGGAAGA 2100
Db |||||

Db 3697 TCGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCCGTATAACCTCAGCACCTCTGGAAGA 3756
QY 2101 C 2101
Db 3757 C 3757
RESULT 3
AAD37234
ID AAD37234 standard; DNA; 3999 BP.
XX
AC AAD37234;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta3990.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 46-47; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus, is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3990 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1 and R2), 8059-10227 (rods R22, R23 and R24, hinge
CC H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids).
XX
SQ Sequence 3999 BP; 1223 A; 907 C; 933 G; 936 T; 0 other;
Query Match 75.2%; Score 1579.8; DB 24; Length 3999;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 1846; Conservative 0; Mismatches 72; Indels 183; Gaps 4;
QY 1 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGAGCCCAT 60
Db |||||
QY 900 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGAGCCCAT 959
Db |||||
QY 61 TCCTTCACAGCATTTGGAGCTCCTGAAGACAAGTCATTGGCAGTTTCATTGATGGAGAG 120
Db |||||
QY 960 TCCTTCACAGCATTTGGAGCTCCTGAAGACAAGTCATTGGCAGTTTCATTGATGGAGAG 1019
Db |||||

QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTCTTTC 180
DB 1020 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTCTTTC 1079
QY 181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
DB 1080 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1139
QY 241 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
DB 1140 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1199
QY 301 TAAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
DB 1200 TAAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1259
QY 361 AACTGAAGTACAAGACGAGATGAATCTCCTAAATTCAAGATGGGAATGCCCTCAGGTTAGC 420
DB 1260 AACTGAAGTACAAGACGAGATGAATCTCCTAAATTCAAGATGGGAATGCCCTCAGGTTAGC 1319
QY 421 TAGCATGGAATAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 480
DB 1320 TAGCATGGAATAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 1379
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAAACAGGAAAATGGAGGAAGA 540
DB 1380 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAAACAGGAAAATGGAGGAAGA 1439
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACATAGGTGCTTCA 600
DB 1440 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACATAGGTGCTTCA 1499
QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTCTCTCACTCACATGGTGGTGTAGT 660
DB 1500 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTCTCTCACTCACATGGTGGTGTAGT 1559
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
DB 1560 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1619
QY 721 AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGCTGGGTCTTTTACAAGACAT 780
DB 1620 AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGCTGGGTCTTTTACAAGACCA 1679
QY 781 CCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTATGTCATGGCTTTCAGA 840
DB 1680 GCCTGAC-----CTAGCTCCTGGACTGACCACCTATTGGAGCCTCTCCTACTCA 1727
QY 841 AAAAGAAGATGCAGTGAACAAGATTACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 900
DB 1728 GACTGTTACTCTGGTGACACA-----ACCTGTGGTTACTTAAGGAAACTGCCATCT- 1777
QY 901 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAACGGGATCTAGAAAAGAAAAAGCAATCCAT 960
DB 1778 ----- 1777
QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGC 1020
DB 1778 -----CCAAACTAGAAATGCCATCTTCTTGTATGTTGGAG----- 1812
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATTAATTAGTCCAAA 1080
DB 1813 ----- 1812
QY 1081 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 1140
DB 1813 -----GTACCTACTCATAGATTACTGCAACAGTTTCCCCCTGGACCTGGA 1856
QY 1141 AAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTAC 1200
DB 1857 AAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTAC 1916
QY 1201 CCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAGAGCTGTATGAACAATGGCA 1260

DB 1917 CCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAAACAATGGCA 1976
QY 1261 AGACCTCCAAGGTGAAATTGAAGTCAACACAGATGTTTATCAAACTTGGATGAAAACAG 1320
DB 1977 AGACCTCCAAGGTGAAATTGAAGTCAACACAGATGTTTATCAAACTTGGATGAAAACAG 2036
QY 1321 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCACTCCTGTTACAAAGACGTTT 1380
DB 2037 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCACTCCTGTTACAAAGACGTTT 2096
QY 1381 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAATAAGTCTCTCAACATTAGGTCCCA 1440
DB 2097 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAATAAGTCTCTCAACATTAGGTCCCA 2156
QY 1441 TTTTGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTTGCAGGAACCTTCTGGT 1500
DB 2157 TTTTGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTTGCAGGAACCTTCTGGT 2216
QY 1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC 1560
DB 2217 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC 2276
QY 1561 AGCAGTTCAAGACAGAAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACATAAGA 1620
DB 2277 AGCAGTTCAAGACAGAAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACATAAGA 2336
QY 1621 ACCTGTATCATGAGTACTCTTGGAGACTGTACGAATATTTTGACAGAGCAGCCCTTTGGA 1680
DB 2337 ACCTGTATCATGAGTACTCTTGGAGACTGTACGAATATTTTGACAGAGCAGCCCTTTGGA 2396
QY 1681 AGGACTAGAGAAACTCTACCAGGAGCCACAGAGAGCTGCCTCCTGAGGAGAGAGCCCCAGAA 1740
DB 2397 AGGACTAGAGAAACTCTACCAGGAGCCACAGAGAGCTGCCTCCTGAGGAGAGAGCCCCAGAA 2456
QY 1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAA 1800
DB 2457 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAA 2516
QY 1801 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 1860
DB 2517 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 2576
QY 1861 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATC 1920
DB 2577 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATC 2636
QY 1921 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 1980
DB 2637 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 2696
QY 1981 GGCACCTCGAGGAGAAAAATTGCGCCTCTGAAAGAGAAACGCTGAGCCAGCTCAATGACCTTGC 2040
DB 2697 GGCACCTCGAGGAGAAAAATTGCGCCTCTGAAAGAGAAACGCTGAGCCAGCTCAATGACCTTGC 2756
QY 2041 TCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTTGGGAAGA 2100
DB 2757 TCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTTGGGAAGA 2816
QY 2101 C 2101
DB 2817 C 2817

RESULT 4
AAD37256
ID AAD37256 standard; DNA; 4966 BP.
XX
AC AAD37256;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-delta3990.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX Chimeric - Homo sapiens.
OS Chimeric - Unidentified.
XX WO200183695-A2.
XX 08-NOV-2001.
PD 27-APR-2001; 2001WO-US13677.
XX 28-APR-2000; 2000US-200777P.
PR (XIAO/) XIAO X.
XX Xiao X;
PI WPI; 2002-049342/06.
XX
XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
XX Example 1; Page 59-60; 71pp; English.
XX
XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence.
XX
SQ Sequence 4966 BP; 1403 A; 1216 C; 1232 G; 1115 T; 0 other;

Query Match 75.2%; Score 1579.8; DB 24; Length 4966;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 1846; Conservative 0; Mismatches 72; Indels 183; Gaps 4;

Qy 1 GAGCTATGCCTACACACAGGCTGCTTATGTCCACCACTCTGACCCCTACACGGAGCCCAATT 60
Db 1657 GAGCTATGCCTACACACAGGCTGCTTATGTCCACCACTCTGACCCCTACACGGAGCCCAATT 1716
Qy 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 120
Db 1717 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 1776
Qy 121 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAAGTATTATCGTGCCTTCTTC 180
Db 1777 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAAGTATTATCGTGCCTTCTTC 1836
Qy 181 TGCTGAGGACACATTGCAAGCAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 240
Db 1837 TGCTGAGGACACATTGCAAGCAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 1896
Qy 241 CCAGTTTCTAFACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 300
Db 1897 CCAGTTTCTAFACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 1956
Qy 301 TAATATCTACAAATTGGGAAGTAAGCTGATTTGGAACAGGAAATTTATCAGAAGATGAAGA 360
Db 1957 TAATATCTACAAATTGGGAAGTAAGCTGATTTGGAACAGGAAATTTATCAGAAGATGAAGA 2016
Qy 361 AACTGAAGTACAGAGCAGATGAATCTCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420

Db 2017 AACTGAAGTACAGAGCAGATGAATCTCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 2076
Qy 421 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480
Db 2077 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 2136
Qy 481 GAAAGAGTTGAATGACTGGCTAAACAAACAGAGAAAGAAACAAGGAAATGGAGGAAGA 540
Db 2137 GAAAGAGTTGAATGACTGGCTAAACAAACAGAGAAAGAAACAAGGAAATGGAGGAAGA 2196
Qy 541 GCCTCTTTGGACCTGATCTTGAAGACCTAAACGCAAGTACAACAACATAAGGTGCTTCA 600
Db 2197 GCCTCTTTGGACCTGATCTTGAAGACCTAAACGCAAGTACAACAACATAAGGTGCTTCA 2256
Qy 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTAGT 660
Db 2257 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTAGT 2316
Qy 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAAACAACCTTAAGGTATTGGG 720
Db 2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAAACAACCTTAAGGTATTGGG 2376
Qy 721 AGATCGATGGGCAACATCTCTAGATGGACAGAGACCGCTGGTCTCTTTTACAAGACAT 780
Db 2377 AGATCGATGGGCAACATCTCTAGATGGACAGAGACCGCTGGTCTCTTTTACAAGACCA 2436
Qy 781 CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTTCA 840
Db 2437 GCCTGAC-----CTAGCTCCTGGACTGACCACTATTGGAGCCTCTCTCTACTCA 2484
Qy 841 AAAAGAAGATGCAGTGAAACAAGATTCACACAACACTGGCTTTAAAGATCAAAATGAATGTT 900
Db 2485 GACTGTCTACTCTGGTGACACA-----ACCTGTGGTTTACTAAGGAAACTGCCATCT- 2534
Qy 901 ATCAAGTCTTCAAAACACTGGCCGCTTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT 960
Db 2535 ----- 2534
Qy 961 GGGCAAACTGTATTCACTCAACAAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db 2535 -----CCAAACTAGAAATGCCATCTTCTCTGTAGTTGGAG----- 2569
Qy 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGTGTGGGATAATTTAGTCCAAAA 1080
Db 2570 ----- 2569
Qy 1081 ACTTGAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA 1140
Db 2570 -----GTACCTACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA 2613
Qy 1141 AAAGTTTCTTGGCTTACAGAAAGCTGAACAACCTGCCAATGTCTCTACAGGATGCTAC 1200
Db 2614 AAAGTTTCTTGGCTTACAGAAAGCTGAACAACCTGCCAATGTCTCTACAGGATGCTAC 2673
Qy 1201 CCGTAAGGAAAGGCTCCTAGAAAGACTCCAAGGAGTAAAGAGCTGATGAACAATGGCA 1260
Db 2674 CCGTAAGGAAAGGCTCCTAGAAAGACTCCAAGGAGTAAAGAGCTGATGAACAATGGCA 2733
Qy 1261 AGACTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCACAACCTGGATGAACAACAG 1320
Db 2734 AGACTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCACAACCTGGATGAACAACAG 2793
Qy 1321 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTTTACAAGACGTTT 1380
Db 2794 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTTTACAAGACGTTT 2853
Qy 1381 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGTCCCA 1440
Db 2854 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGTCCCA 2913
Qy 1441 TTTGGAAGCCAGTTCTGACCCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGT 1500

Db 2914 TTTGGAAGCCAGTTCTTGACCACTGGAAGCGTCTGCACTTTCTCTGCAGGAACCTTCTGGT 2973

2y 1501 GTGGCTACAGCTGAAAGATGATGAATTAAAGCCGGCAGGCACCTATTGGAGGCGACTTTC 1560

Db 2974 GTGGCTACAGCTGAAAGATGATGAATTAAAGCCGGCAGGCACCTATTGGAGGCGACTTTC 3033

2y 1561 AGCAGTTGAGAAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACTAAAGA 1620

Db 3034 AGCAGTTGAGAAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACTAAAGA 3093

2y 1621 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTTGACAGAGCAGCCTTTGGA 1680

Db 3094 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTTGACAGAGCAGCCTTTGGA 3153

2y 1681 AGGACTAGAGAAACTCTACCAGAGCCAGAGAGCTGCCCTCCTGAGGAGAGAGCCAGAA 1740

Db 3154 AGGACTAGAGAAACTCTACCAGAGCCAGAGAGCTGCCCTCCTGAGGAGAGAGCCAGAA 3213

2y 1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAA 1800

Db 3214 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAA 3273

Qy 1801 CCTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGAACT 1860

Db 3274 CCTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGAACT 3333

2y 1861 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATC 1920

Db 3334 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATC 3393

2y 1921 CTGGCAGCCCGTGGCGGATCTCCTCAATTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAA 1980

Db 3394 CTGGCAGCCCGTGGCGGATCTCCTCAATTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAA 3453

2y 1981 GGCACCTCGAGAGAAATTCGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 2040

Db 3454 GGCACCTCGAGAGAAATTCGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 3513

Qy 2041 TCGCCAGCTTACCACTTTGGGCATTACAGCTCTCACCGTATACCTCAGCACTCTGGGAAGA 2100

Db 3514 TCGCCAGCTTACCACTTTGGGCATTACAGCTCTCACCGTATACCTCAGCACTCTGGGAAGA 3573

Qy 2101 C 2101

Db 3574 C 3574

RESULT 5

AAD37262

ID AAD37262 standard; DNA; 4990 BP.

XX

AC AAD37262;

XX

DT 21-AUG-2002 (first entry)

XX

DE Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3990.

XX

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

KW Becker muscular dystrophy; ds.

XX

OS Chimeric - Homo sapiens.

OS Chimeric - Cytomegalovirus.

OS Chimeric - Unidentified.

XX

PN WO200183695-A2.

XX

PD 08-NOV-2001.

XX

PF 27-APR-2001; 2001WO-US13677.

XX

PR 28-APR-2000; 2000US-200777P.

XX

PA (XIAO/) XIAO X.

XX

PI Xiao X;

XX

DR WPI; 2002-049342/06.

XX

PT New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,

PT rod repeats, H1 and H4 domains and a cysteine rich domain of a

PT dystrophin gene -

XX

PS Example 1; Page 67-68; 71pp; English.

XX

CC The present invention relates to an isolated nucleotide sequence encoding

CC a dystrophin minigene. The minigene comprises N-terminal or modified

CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4

CC domains and cysteine-rich domains of dystrophin or utrophin genes. The

CC invention also relates to a recombinant adeno-associated virus (AAV)

CC comprising dystrophin minigene operably linked to an expression control

CC element. The dystrophin minigene in operable linkage with an expression

CC control element, in a recombinant adeno-associated virus or retrovirus is

CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular

CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV

CC vector plasmid construct containing human dystrophin minigenes, a

CC cytomagalovirus (CMV) promoter and a small polyA signal sequence.

XX

SQ Sequence 4990 BP; 1439 A; 1185 C; 1208 G; 1158 T; 0 other;

Qy Query Match 75.2%; Score 1579.8; DB 24; Length 4990;

Db Best Local Similarity 87.9%; Pred. No. 0;

Qy Matches 1846; Conservative 0; Mismatches 72; Indels 183; Gaps 4;

1 GAGCTATGCTTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAAT 60

1681 GAGCTATGCTTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAAT 1740

61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCAATTTGSCAGTTCATTGATGGAGAG 120

1741 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCAATTTGSCAGTTCATTGATGGAGAG 1800

121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATTTATCGTGGCTTCTTTC 180

1801 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATTTATCGTGGCTTCTTTC 1860

181 TGCTGAGGACACATTGCAAGCAACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240

1861 TGCTGAGGACACATTGCAAGCAACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1920

241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGCCGGGTTGG 300

1921 CCAGTTTCATACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGCCGGGTTGG 1980

301 TAATATTCTACAATTGGGAAGTAAGCTGANTGGAACAGGAAATTTATCAGAAGATGAAGA 360

1981 TAATATTCTACAATTGGGAAGTAAGCTGANTGGAACAGGAAATTTATCAGAAGATGAAGA 2040

361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420

2041 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 2100

421 TAGCATGGAATAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480

2101 TAGCATGGAATAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 2160

481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAACAGGAAATGGAGGAAGA 540

2161 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAACAGGAAATGGAGGAAGA 2220

541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACACATAAGGTGCTTCA 600

2221 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACACATAAGGTGCTTCA 2280

601 AGAAGATCTAGAACACAGAAACAAGTCAGGGGTCAATTCTCTCACTCACATGGTGGTAGT 660

Db 2281 AGAAGACTAGAAACAAGAACAGTCAGGGTCAATTCTCTCACTCACATGCTGGTGTAGT 2340
Qy 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
Db 2341 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 2400
Qy 721 AGATCGATGGGCAACATCTGTAGATGGACAGAACGCGTGGTTCCTTTACAAGACAT 780
Db 2401 AGATCGATGGGCAACATCTGTAGATGGACAGAACGCGTGGTTCCTTTACAAGACCA 2460
Qy 781 CCTTCTCAAAATGGCAACGCTTCTACTGAAGAACAGTGCCTTTTAGTGCAATGGCTTTCAGA 840
Db 2461 GCCTGAC-----CTAGCTCCTGGACTGACCACTATTGGAGCCTCTCCTACTCA 2508
Qy 841 AAAAGAAAGATGCAGTGAACAAGATTACACAACTGGCTTTAAAGATCAAAATGAAATGTT 900
Db 2509 GACTGTTACTCTGGTGACACA-----ACCTGTGGTTACTAAGGAAACTGCCATCT- 2558
Qy 901 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAAGAAAAGCAATCCAT 960
Db 2559 ----- 2558
Qy 961 GGGCAACTGTATTCACTCAAAACAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db 2559 -----CCAAACTAGAAATGCCATCTTCTCCTGTGATGTTGGAG----- 2593
Qy 1021 CCAGAAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTTAGTCCAAAA 1080
Db 2594 ----- 2593
Qy 1081 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTTCCCCCTGGACCTGGA 1140
Db 2594 -----GTACCTACTCATAGATTACTGCAACAGTTTCCCCCTGGACCTGGA 2637
Qy 1141 AAAGTTTCTTGCTGGCTTACAGAACTGAAACAACTGCCAATGTCTCTACAGGATGCTAC 1200
Db 2638 AAAGTTTCTTGCTGGCTTACAGAACTGAAACAACTGCCAATGTCTCTACAGGATGCTAC 2697
Qy 1201 CCGTAAGGAAAGCTCCTAGAACTCCAAAGGGAGTAAAGAGCTGATGAAACAATGGCA 1260
Db 2698 CCGTAAGGAAAGCTCCTAGAACTCCAAAGGGAGTAAAGAGCTGATGAAACAATGGCA 2757
Qy 1261 AGACCTCCAAGTGAAATGAAGCTCACACAGATGTTTTATCAAACTGGATGAAACAG 1320
Db 2758 AGACCTCCAAGTGAAATGAAGCTCACACAGATGTTTTATCAAACTGGATGAAACAG 2817
Qy 1321 CCAAAAAATCCTGAGATCCCTGGAAGTTCCGATGATGCAGTCTCTGTTTACAAAGACGTTT 1380
Db 2818 CCAAAAAATCCTGAGATCCCTGGAAGTTCCGATGATGCAGTCTCTGTTTACAAAGACGTTT 2877
Qy 1381 GGATAACATGAACCTCAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAGGTCCCA 1440
Db 2878 GGATAACATGAACCTCAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAGGTCCCA 2937
Qy 1441 TTTGGAAGCCAGTTCTGACCACTGGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGT 1500
Db 2938 TTTGGAAGCCAGTTCTGACCACTGGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGT 2997
Qy 1501 GTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC 1560
Db 2998 GTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC 3057
Qy 1561 ACCAGTTTCAGAAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGGAATTGAAAACTAAAGA 1620
Db 3058 ACCAGTTTCAGAAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGGAATTGAAAACTAAAGA 3117
Qy 1621 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGA 1680
Db 3118 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGA 3177
Qy 1681 AGGACTAGAGAAACTCTACAGGAGCCCGAGAGTGCCTCCTGAGGAGAGAGCCCGAGAA 1740

Db 3178 AGGACTAGAGAAACTCTACAGGAGCCAGAGAGCTGCCTCTGAGGAGAGAGCCCGAGAA 3237
Qy 1741 TGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAA 1800
Db 3238 TGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAA 3297
Qy 1801 CCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 1860
Db 3298 CCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 3357
Qy 1861 TCAAGAGGCCACGGATGAGTGGACCTCAAGCTTGCCTCAAGATCACCTCGAGAAAAGTCAA 1920
Db 3358 TCAAGAGGCCACGGATGAGTGGACCTCAAGCTTGCCTCAAGATCACCTCGAGAAAAGTCAA 3417
Qy 1921 CTGGCAGCCCGTGGCGATCTCTCATTTGACTCTCTTCCAAAGATCACCTCGAGAAAAGTCAA 1980
Db 3418 CTGGCAGCCCGTGGCGATCTCTCATTTGACTCTCTTCCAAAGATCACCTCGAGAAAAGTCAA 3477
Qy 1981 GGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 2040
Db 3478 GGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 3537
Qy 2041 TCGCCAGCTTACCACCTTTGGCGATTGAGCTCTCACCCGATTAACCTCAGCACCTCTGGAAGA 2100
Db 3538 TCGCCAGCTTACCACCTTTGGCGATTGAGCTCTCACCCGATTAACCTCAGCACCTCTGGAAGA 3597
Qy 2101 C 2101
Db 3598 C 3598

RESULT 6

AAD37237
ID AAD37237 standard; DNA; 3858 BP.

XX
AC AAD37237;

XX
DT 21-AUG-2002 (first entry)

XX
DE Human dystrophin minigene delta3849.

XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

XX
KW Becker muscular dystrophy; ds.

XX
OS Homo sapiens.

XX
PN WO200183695-A2.

XX
PD 08-NOV-2001.

XX
PF 27-APR-2001; 2001WO-US13677.

XX
PR 28-APR-2000; 2000US-200777P.

XX
PA (XIAO/) XIAO X.

XX
PI Xiao X;

XX
DR WPI; 2002-049342/06.

XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
dystrophy comprises an N-terminal domain or modified N-terminal domain,
rod repeats, H1 and H4 domains and a cysteine rich domain of a
dystrophin gene -

PS
XX Example 1; Page 48-49; 71pp; English.

XX
CC The present invention relates to an isolated nucleotide sequence encoding
a dystrophin minigene. The minigene comprises N-terminal or modified
N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
domains and cysteine-rich domains of dystrophin or utrophin genes. The
invention also relates to a recombinant adeno-associated virus (AAV)

CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3849 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1, R2), 8059-10227 (rods R22, R23 and R24, hinge H4
CC and CR domain) and 11047-11058 (dystrophin last 3 amino acids).
XX
SQ Sequence 3858 BP; 1189 A; 866 C; 905 G; 898 T; 0 other;

Query Match 68.7%; Score 1443; DB 24; Length 3858;
Best Local Similarity 84.6%; Pred. No. 0;
Matches 1777; Conservative 0; Mismatches 0; Indels 324; Gaps 1;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGGAGCCCAT 60
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 900 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGGAGCCCAT 959
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 120
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 960 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1019
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATATCGTGGCTTCTTTC 180
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1020 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATATCGTGGCTTCTTTC 1079
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 181 TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1080 TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1139
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 241 CCAGTTTTCATACACTGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1140 CCAGTTTTCATACACTGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1199
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 301 TAATATTCTACRAATGGGAAGTAAGCTGATTGGAACAGGAAGAAATATCAGAAGATGAAGA 360
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1200 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGAAGAAATATCAGAAGATGAAGA 1259
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 361 AACTGAAGTACAAGACAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1260 AACTGAAGTACAAGACAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1319
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1320 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 1379
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAGGAAAAATGGAGGAAGA 540
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1380 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAGGAAAAATGGAGGAAGA 1439
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAAACAATAAGGTGCTTCA 600
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1440 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAAACAATAAGGTGCTTCA 1499
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGTCAATTTCTCACTACATGGTGGTAGT 660
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1500 AGAAGATCTAGAACAAAGAACAAAGTCAGGTCAATTTCTCACTACATGGTGGTAGT 1559
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1560 TGATGAATCTAGTGGAGATCACGCAACTGTGCTTTGGAAGAACAACTTAAGGTATTGGG 1619
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 721 AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGTGGGTTCTTTTACAAGACAT 780
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1620 AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGTGGGTTCTTTTACAAGAC-- 1677
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 781 CCTTCTCAATGGCAACCTCTTACTGAAGAACAGTGCCTTTTATGTCATGGCTTTCAGA 840
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1678 ----- 1677
QY 841 AAAAGAAGATGCAGTGAACAAGATTTCACACACTGGCTTTAAAGATCAAAATGAAATGTT 900
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 1678 ----- 1677
QY 901 ATCAAGTCTTCAAAAACTGGCCGTTTAAAGCGGATCTAGAAAAAGCAATCCAT 960
Db 1678 ----- 1677
QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db 1678 ----- 1677
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
Db 1678 ----- 1677
QY 1081 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 1140
Db 1678 -----ACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 1715
QY 1141 AAAAGTTTCTTGCTGGCTTACAGAACTGAAACAACTGCCAATGTCTCTACAGGATGCTAC 1200
Db 1716 AAAAGTTTCTTGCTGGCTTACAGAACTGAAACAACTGCCAATGTCTCTACAGGATGCTAC 1775
QY 1201 CCGTAAGGAAAGGCTCCTAGAAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 1260
Db 1776 CCGTAAGGAAAGGCTCCTAGAAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 1835
QY 1261 AGACCTCCAAGTGAAATTTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACAG 1320
Db 1836 AGACCTCCAAGTGAAATTTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACAG 1895
QY 1321 CCAAAAAATCCTGAGATCCCTGGAAGTTCCGATGATGCAGTCTGTTACAAAGACGTTT 1380
Db 1896 CCAAAAAATCCTGAGATCCCTGGAAGTTCCGATGATGCAGTCTGTTACAAAGACGTTT 1955
QY 1381 GGATAACATGAACTTCAAGTGGAGTGAATTCGGAAAAAGTCTCTCAACATTAGGTCCCA 1440
Db 1956 GGATAACATGAACTTCAAGTGGAGTGAATTCGGAAAAAGTCTCTCAACATTAGGTCCCA 2015
QY 1441 TTTGGAAAGCCAGTTCTGACCAGTGGAAAGCTGTGCACCTTTCTCTGCAGGAACCTTCTGGT 1500
Db 2016 TTTGGAAAGCCAGTTCTGACCAGTGGAAAGCTGTGCACCTTTCTCTGCAGGAACCTTCTGGT 2075
QY 1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGCAGCTTTCC 1560
Db 2076 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGCAGCTTTCC 2135
QY 1561 AGCAGTTCAGAACAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAAAACTAAAGA 1620
Db 2136 AGCAGTTCAGAACAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAAAACTAAAGA 2195
QY 1621 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGA 1680
Db 2196 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGA 2255
QY 1681 AGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCGAGAA 1740
Db 2256 AGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCGAGAA 2315
QY 1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACCTAGTGGGAAAAATTGAA 1800
Db 2316 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACCTAGTGGGAAAAATTGAA 2375
QY 1801 CCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAAC 1860
Db 2376 CCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAAC 2435
QY 1861 TCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATC 1920
Db 2436 TCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATC 2495
QY 1921 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACTCGAGAAAGTCAA 1980
Db 2496 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACTCGAGAAAGTCAA 2555

Db 1920 CCAGTTTCATACTCATGAGGGTACATGATGGATTGTGACAGCCCATCAGGGCCGGTTGG 1979

QY 301 TAATAATCTACAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAGATGAAGA 360

Db 1980 TAATAATCTACAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAGATGAAGA 2039

QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTTAATTTCAAGATGGGAATGCCTCAGGGTAGC 420

Db 2040 AACTGAAGTACAAGAGCAGATGAATCTCTTAATTTCAAGATGGGAATGCCTCAGGGTAGC 2099

QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 480

Db 2100 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 2159

QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGAAATGGAGGAAGA 540

Db 2160 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGAAATGGAGGAAGA 2219

QY 541 GCCTCTTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAACAACATAAAGGTGCTCA 600

Db 2220 GCCTCTTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAACAACATAAAGGTGCTCA 2279

QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGTCATTTCTCTCACTCACATGGTGGTGTAGT 660

Db 2280 AGAAGATCTAGAACAAAGAACAAAGTCAGGTCATTTCTCTCACTCACATGGTGGTGTAGT 2339

QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAACAACTTAAGGTATTGGG 720

Db 2340 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAACAACTTAAGGTATTGGG 2399

QY 721 AGATCGATGGGAAACATCTGTAGATGGACAGAAACCGCTGGTCTCTTTTACAAGACAT 780

Db 2400 AGATCGATGGGAAACATCTGTAGATGGACAGAAACCGCTGGTCTCTTTTACAAGAC-- 2457

QY 781 CTTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTATGTCATGGCTTTCAGA 840

Db 2458 ----- 2457

QY 841 AAAAGAAGATGCAGTGAACAAGATTCAACAACCTGGCTTTAAAGATCAAAATGAAATGTT 900

Db 2458 ----- 2457

QY 901 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960

Db 2458 ----- 2457

QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020

Db 2458 ----- 2457

QY 1021 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1080

Db 2458 ----- 2457

QY 1081 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCTCGACCTGGA 1140

Db 2458 -----ACTCATAGATTACTGCAACAGTTCCCTCGACCTGGA 2495

QY 1141 AAAGTTTCTTGCTGGCTTACAGAAAGCTGAAACAACTGCCAATGTCTCCTACAGGATGCTAC 1200

Db 2496 AAAGTTTCTTGCTGGCTTACAGAAAGCTGAAACAACTGCCAATGTCTCCTACAGGATGCTAC 2555

QY 1201 CCGTAAGGAAAGGCTCCTAGAAAGACTCCAAGGGAGTAAAGAGCTGATGAACAAATGGCA 1260

Db 2556 CCGTAAGGAAAGGCTCCTAGAAAGACTCCAAGGGAGTAAAGAGCTGATGAACAAATGGCA 2615

QY 1261 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAG 1320

Db 2616 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAG 2675

QY 1321 CCAAAAAATCCTGAGATCCCTGGAGGTTCCGATGATGACGTCCTGTTTACAAAGACGTTT 1380

Db 2676 CCAAAAAATCCTGAGATCCCTGGAGGTTCCGATGATGACGTCCTGTTTACAAAGACGTTT 2735

QY 1381 GGATAACATGAACCTTCAAGTGGAGTGAACTTCGGAAGAAAGTCTCTCAACATTAGGTCCCA 1440

Db 2736 GGATAACATGAACCTTCAAGTGGAGTGAACTTCGGAAGAAAGTCTCTCAACATTAGGTCCCA 2795

QY 1441 TTTTGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACTTTCTCTGCAGGAACCTTCTGGT 1500

Db 2796 TTTTGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACTTTCTCTGCAGGAACCTTCTGGT 2855

QY 1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTTATTGGAGGGGACCTTCC 1560

Db 2856 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTTATTGGAGGGGACCTTCC 2915

QY 1561 AGCAGTTTCAAGAACAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAAAACTAAAGA 1620

Db 2916 AGCAGTTTCAAGAACAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAAAACTAAAGA 2975

QY 1621 ACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGA 1680

Db 2976 ACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGA 3035

QY 1681 AGGACTAGAGAAAACTCTACCAGAGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCCGAGAA 1740

Db 3036 AGGACTAGAGAAAACTCTACCAGAGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCCGAGAA 3095

QY 1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAA 1800

Db 3096 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAA 3155

QY 1801 CCTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAGAGACTCCAGGAAT 1860

Db 3156 CCTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAGAGACTCCAGGAAT 3215

QY 1861 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATC 1920

Db 3216 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATC 3275

QY 1921 CTGGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 1980

Db 3276 CTGGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 3335

QY 1981 GGCACCTTCGAGGAGAAAAATTGGCCCTCTGAAAGAGAACCTGAGCCACGTCAATGACCTTGC 2040

Db 3336 GGCACCTTCGAGGAGAAAAATTGGCCCTCTGAAAGAGAACCTGAGCCACGTCAATGACCTTGC 3395

QY 2041 TCGCCAGCTTACCACTTTGGGCAATTCAGTCTCACCGTATACCTCAGCACCTCTGGAAGA 2100

Db 3396 TCGCCAGCTTACCACTTTGGGCAATTCAGTCTCACCGTATACCTCAGCACCTCTGGAAGA 3455

QY 2101 C 2101

Db 3456 C 3456

RESULT 9

AAD37264

ID AAD37264 standard; DNA; 5060 BP.

XX

AC AAD37264;

XX

DT 21-AUG-2002 (first entry)

XX

DE Adeno-associated virus (AAV) vector plasmid, AAV-E-CMV-3849.

XX

KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

XX

KW Becker muscular dystrophy; ds.

XX

OS Chimeric - Homo sapiens.

OS

OS Chimeric - Cytomegalovirus.

OS

XX

OS Chimeric - Unidentified.

PN WO200183695-A2.

08-NOV-2001. 1
27-APR-2001; 2001WO-US13677. 1
28-APR-2000; 2000US-200777P. 1
(XIAO/) XIAO X. 1
Xiao X; 1
WPI; 2002-049342/06. 1
New dystrophin minigene for treating Duchenne or Becker muscular dystrophy comprises an N-terminal domain or modified N-terminal domain, rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene - 1
Example 1; Page 70-71; 71pp; English. 1
The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is AAV vector plasmid construct containing human dystrophin minigenes, a muscle creatine kinase (MCK) enhancer, a cytomegalovirus (CMV) promoter and a small polyA signal sequence. 1
Sequence 5060 BP; 1449 A; 1217 C; 1234 G; 1160 T; 0 other; 1
Query Match 68.7%; Score 1443; DB 24; Length 5060; 1
Best Local Similarity 84.6%; Pred. No. 0; 1
Matches 1777; Conservative 0; Mismatches 0; Indels 324; Gaps 1; 1
1 GAGCTATGCTTACACACAGGCTGCTTATGTACCACTCTGACCCCTACACGGAGCCATT 60 1
1892 GAGCTATGCTTACACACAGGCTGCTTATGTACCACTCTGACCCCTACACGGAGCCATT 1951 1
61 TCCTTCACAGCATTTGGAAGCTCTGAAAGCAAGTCAATTTGGCAGTTTCAATGATGGAGAG 120 1
1952 TCCTTCACAGCATTTGGAAGCTCTGAAAGCAAGTCAATTTGGCAGTTTCAATGATGGAGAG 2011 1
121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATATCGTGGCTTCTTTC 180 1
2012 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATATCGTGGCTTCTTTC 2071 1
181 TGCTGAGGACACATTTGGAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240 1
2072 TGCTGAGGACACATTTGGAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 2131 1
241 CCAGTTTCATACATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG 300 1
2132 CCAGTTTCATACATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG 2191 1
301 TAATATTCTACAAATTTGGGAAGTAAGCTGATTTGGAACAGGAATATTATCAGAAGATGAAGA 360 1
2192 TAATATTCTACAAATTTGGGAAGTAAGCTGATTTGGAACAGGAATATTATCAGAAGATGAAGA 2251 1
361 AACTGAAGTACAGAGCAGATGAATCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420 1
2252 AACTGAAGTACAGAGCAGATGAATCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 2311 1
421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480 1
2312 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 2371 1
481 GAAAGAGTTGAATGACTGGCTTAAACAAAACAGAAAGAAACAGGAAAAATGGAGGAAGA 540 1

Db 2372 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAACAGGAAAATGGAGGAAGA 2431 1
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACATAGGTGCTTCA 600 1
Db 2432 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACATAGGTGCTTCA 2491 1
QY 601 AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATCTCTCTCACTCACTCATGGTGGTAGT 660 1
Db 2492 AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATCTCTCTCACTCACTCATGGTGGTAGT 2551 1
QY 661 TGATGAATCTAGTGGAGATCACCGCAACTGCTGCTTTTGAAGAAACAACTTAAAGTATTGGG 720 1
Db 2552 TGATGAATCTAGTGGAGATCACCGCAACTGCTGCTTTTGAAGAAACAACTTAAAGTATTGGG 2611 1
QY 721 AGATCGATGGCAACACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTCAAGACAT 780 1
Db 2612 AGATCGATGGCAACACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTCAAGAC-- 2669 1
QY 781 CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTTAGTGATGGCTTTTCA 840 1
Db 2670 ----- 2669 1
QY 841 AAAAGAGATGCAGTGAACAAGATTACACAACACTGGCTTTAAAGATCAAAATGAATGTT 900 1
Db 2670 ----- 2669 1
QY 901 ATCAAGTCTTCAAAAACCTGGCCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960 1
Db 2670 ----- 2669 1
QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCACTGAC 1020 1
Db 2670 ----- 2669 1
QY 1021 CCAGAAAGCAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1080 1
Db 2670 ----- 2669 1
QY 1081 ACTTGAAGAGATACAGCACAGACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 1140 1
Db 2670 -----ACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 2707 1
QY 1141 AAAGTTTCTTGGCTTACAGAAAGTGAACAACTGCCAATGTCTACAGGATGCTAC 1200 1
Db 2708 AAAGTTTCTTGGCTTACAGAAAGTGAACAACTGCCAATGTCTACAGGATGCTAC 2767 1
QY 1201 CCGTAAGGAAAGGCTCCTAGAAAGTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 1260 1
Db 2768 CCGTAAGGAAAGGCTCCTAGAAAGTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 2827 1
QY 1261 AGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAG 1320 1
Db 2828 AGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAG 2887 1
QY 1321 CCAAAAAATCCTGAGATCCCTGGAAGTTCGGATGATGCGAGTCTGTACAAAGACGTTT 1380 1
Db 2888 CCAAAAAATCCTGAGATCCCTGGAAGTTCGGATGATGCGAGTCTGTACAAAGACGTTT 2947 1
QY 1381 GGATAACATGAACCTCAAGTGGAGTGAACCTCGGAAAAGTCTCTCAACATTAGGTCCCA 1440 1
Db 2948 GGATAACATGAACCTCAAGTGGAGTGAACCTCGGAAAAGTCTCTCAACATTAGGTCCCA 3007 1
QY 1441 TTTGGAAGCCAGTTCTGACCAGTGGAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGTT 1500 1
Db 3008 TTTGGAAGCCAGTTCTGACCAGTGGAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGTT 3067 1
QY 1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTTATTTGGAGGCGACTTCC 1560 1
Db 3068 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTTATTTGGAGGCGACTTCC 3127 1
QY 1561 AGCAGTTCAAGACAGAACGATGTACATAGGGCTTCAAGAGGGAATTTGAAACTAAAGA 1620 1

Db 3128 AGCAGTTGAGAACGATGTTACATAGGGCCCTCAAGAGGGAATTGAAACTAAAGA 3187
QY 1621 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGA 1680
Db 3188 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGA 3247
QY 1681 AGGACTAGAGAAACTCTACGAGAGCCAGAGAGCTGCTCCTGAGGAGAGAGCCAGAA 1740
Db 3248 AGGACTAGAGAAACTCTACGAGAGCCAGAGAGCTGCTCCTGAGGAGAGAGCCAGAA 3307
QY 1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATFACTGAGTGGGAAATTTGAA 1800
Db 3308 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATFACTGAGTGGGAAATTTGAA 3367
QY 1801 CCTGCACCTCGCTGACTGGCAGAGAGAAATATAGATGAGACCCCTTGAAAGACTCCAGGAACT 1860
Db 3368 CCTGCACCTCGCTGACTGGCAGAGAGAAATATAGATGAGACCCCTTGAAAGACTCCAGGAACT 3427
QY 1861 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATC 1920
Db 3428 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATC 3487
QY 1921 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 1980
Db 3488 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 3547
QY 1981 GGCACCTTCGAGGAGAAATTCGCCCTCTGAAAGAGAACGCTGAGCCACGTCATGACCTTGC 2040
Db 3548 GGCACCTTCGAGGAGAAATTCGCCCTCTGAAAGAGAACGCTGAGCCACGTCATGACCTTGC 3607
QY 2041 TCGCCAGCTTACACACTTTGGGCATTCAGCTCTCACCCGTATTAACCTCAGCACCTCTGGAAGA 2100
Db 3608 TCGCCAGCTTACACACTTTGGGCATTCAGCTCTCACCCGTATTAACCTCAGCACCTCTGGAAGA 3667
QY 2101 C 2101
Db 3668 C 3668

RESULT 10
ABK82000
ID ABK82000 standard; DNA; 8689 BP.
XX
AC ABK82000;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding mini-dystrophin protein deltaH2-R19.
XX
KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200229056-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US31126.
XX
PR 06-OCT-2000; 2000US-238848P.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Chamberlain JS, Harper SQ;
XX
DR WPI; 2002-435334/46.
XX

PT A composition for preparing therapeutic drugs, has a mini-dystrophin
PT peptide comprising a specific number of spectrin-like repeat domains,
PT or a nucleic acid sequence encoding the mini-dystrophin peptide -
XX

PS Disclosure; Fig 15; 145pp; English.
XX
CC The invention describes a composition comprising a mini-dystrophin
CC peptide comprising a spectrin-like repeat domain, where the domain
CC comprises n spectrin-like repeats, and contains no more than n
CC spectrin-like repeats, where n is an even number between 4-24, or a
CC nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin
CC peptide or the polynucleotide encoding it is useful as a medicament,
CC for preparing a drug for therapeutic application and in the preparation
CC of a composition for treatment of muscle disease, e.g. Duchenne's
CC muscular dystrophy (DMD). This sequence represents a mini-dystrophin
CC sequence of the invention.

XX
SQ Sequence 8689 BP; 2721 A; 1804 C; 1861 G; 2303 T; 0 other;
Query Match 61.5%; Score 1293; DB 24; Length 8689;
Best Local Similarity 72.5%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 798; Gaps 1;

QY 1 GAGCTATGCTACACACAGGCTGCTTATGTACACCCTGTGACCCCTACACGGAGCCCAT 60
Db 1099 GAGCTATGCTACACACAGGCTGCTTATGTACACCCTGTGACCCCTACACGGAGCCCAT 1158
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
Db 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAGTATTCGTGGCTTCTTTC 180
Db 1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAGTATTCGTGGCTTCTTTC 1278
QY 181 TGCTGAGGACACATTCGAAGCACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1279 TGCTGAGGACACATTCGAAGCACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
QY 241 CCAGTTTCATCTACTGAGGGGTACATGATGATTTTGACAGCCCATCAGGGCCGGTTGG 300
Db 1339 CCAGTTTCATCTACTGAGGGGTACATGATGATTTTGACAGCCCATCAGGGCCGGTTGG 1398
QY 301 TAATATTCTAATTTGGGAAGTAAGCTGATTGGAACAGGAAAATATCAGAAGATGAAGA 360
Db 1399 TAATATTCTAATTTGGGAAGTAAGCTGATTGGAACAGGAAAATATCAGAAGATGAAGA 1458
QY 361 AACTGAAGTACAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1518
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGATTTTAAATGGATCTCCAGAATCAGAAACT 480
Db 1519 TAGCATGGAAAAACAAAGCAATTTACATAGATTTTAAATGGATCTCCAGAATCAGAAACT 1578
QY 481 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAACAGGAAAATGGAGGAAGA 540
Db 1579 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAACAGGAAAATGGAGGAAGA 1638
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCCAAGTACAACAATAAGGTGCTTCA 600
Db 1639 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCCAAGTACAACAATAAGGTGCTTCA 1698
QY 601 AGAAGATCTAGAACACAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTAGT 660
Db 1699 AGAAGATCTAGAACACAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTAGT 1758
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 720
Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 1818
QY 721 AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 780
Db 1819 AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 1878
QY 781 CCTTCTCAAATGGCAACGTTCTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 840

Db 1879 CCTTCTCAAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTTAGTGCAATGGCTTTTCAGA 1938
Qy 841 AAAAGAAGATGCAGTGAAACAAGATTTCACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT 900
Db 1939 AAAAGAAGATGCAGTGAAACAAGATTTCACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT 1998
Qy 901 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
Db 1999 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2058
Qy 961 GGGCAAACTGTATTCTACTCAAAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db 2059 GGGCAAACTGTATTCTACTCAAAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGAC 2118
Qy 1021 CCAGAAGACGGAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAAATTAGTCCAAA 1080
Db 2119 CCAGAAGACGGAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAAATTAGTCCAAA 2178
Qy 1081 ACTTGAAAAGAGTACAGCACAG----- 1102
Db 2179 ACTTGAAAAGAGTACAGCACAGATTTTCACAGCAGCCTGACCTAGCTCCTGGACTGACCAC 2238
Qy 1103 ----- 1102
Db 2239 TATTGGAGCCTCTCCTACTCAGACTGTACTCTGTGTGACACAACTGTGTACTAAGGA 2298
Qy 1103 ----- 1102
Db 2299 AACTGCCATCTCCAAAACTAGAAATGCCATCTCCTTGATGTGAGGTACCTGTCTGGC 2358
Qy 1103 ----- 1102
Db 2359 AGATTTCAACCGGCTTGGACAGAACTTACCAGACTGGCTTTCTCTGCTTGATCAAGTTAT 2418
Qy 1103 ----- 1102
Db 2419 AAAATCACAGAGGGTGATGGTGGTGACCTTGAGGATATCAACGAGATGATCATCAAGCA 2478
Qy 1103 ----- 1102
Db 2479 GAAGGCAACAATGCAGGATTTGGAACAGAGCGGTCCCAGTTGGAAGAACTCATTTACCGC 2538
Qy 1103 ----- 1102
Db 2539 TGCCCCAAAATTTGAAAAACAAGACCAGCAATCAAGAGGCTAGAACAAATCATACGGATCG 2598
Qy 1103 ----- 1102
Db 2599 AATTGAAAGAATTCAGAATCAGTGGGATGAAGTACAAGAACACCTTCAGAACCGGAGGCA 2658
Qy 1103 ----- 1102
Db 2659 ACAGTTGAATGAAATGTTAAAGGATTCACACAATGGCTGGAAGCTAAGGAAGAAGCTGA 2718
Qy 1103 ----- 1102
Db 2719 GCAGGTCTTAGGACAGGCCAGAGCCAAGCTTGAGTCAATGGAAGGAGGTCCCTATACAGT 2778
Qy 1103 ----- 1102
Db 2779 AGATGCAATCCAAAAGAAAATCACAGAAACCAAGCAGTTGGCCAAAGACCTCCGCCAGTG 2838
Qy 1103 ----- 1102
Db 2839 GCAGACAAATGTAGATGTGGCAAAATGACTTTGGSCCTGAAACTTCTCCGGGATTATTCTGC 2898
Qy 1103 ----- 1102
Db 2899 AGATGATACCAGAAAAGTCCACATGATAACAGAGAATATCAATGCCCTTTGGAGAAAGCAT 2958
Qy 1103 ----- 1122
Db 2959 TCATAAAAGGGTGAGTGACGAGAGGCTGCTTTGGGAAGAACTCATAGATTACTGCAACA 3018

RESULT 11
AAD06794
ID AAD06794 standard; DNA; 5952 BP.

Qy 1123 GTTCCCCCTGGACCTGGAAAAGTTTCTTGGCTGGCTTACAGAAAGCTGAAAACAACTGCCAA 1182
Db 3019 GTTCCCCCTGGACCTGGAAAAGTTTCTTGGCTGGCTTACAGAAAGCTGAAAACAACTGCCAA 3078
Qy 1183 TGTCTTACAGATGCTACCCGTAAGGAAAGGCTCCTAGAAAGCTCCAAGGGAGTAAAGA 1242
Db 3079 TGTCTTACAGATGCTACCCGTAAGGAAAGGCTCCTAGAAAGCTCCAAGGGAGTAAAGA 3138
Qy 1243 GCTGATGAAACAATGGCAAGACCTCCAAGGTGAAAATTGAAGCTCACACAGATGTTTATCA 1302
Db 3139 GCTGATGAAACAATGGCAAGACCTCCAAGGTGAAAATTGAAGCTCACACAGATGTTTATCA 3198
Qy 1303 CAACCTGGATGAAAAACAGCCAAAATCCTGAGATCCCTGGAAAGTTCGGAAAAAGTC 1362
Db 3199 CAACCTGGATGAAAAACAGCCAAAATCCTGAGATCCCTGGAAAGTTCGGAAAAAGTC 3258
Qy 1363 CCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTC 1422
Db 3259 CCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTC 3318
Qy 1423 TCTCAACATTAGGTCCCATTGGAAGCCAGTTTCTGACCAGTGGAAAGCTCTGCACCTTTC 1482
Db 3319 TCTCAACATTAGGTCCCATTGGAAGCCAGTTTCTGACCAGTGGAAAGCTCTGCACCTTTC 3378
Qy 1483 TCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCAAC 1542
Db 3379 TCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCAAC 3438
Qy 1543 TATTGGAGGCGACTTTCAGCAGTTCCAGAACGATGATGAATTAAGCCGGCAGGCAAC 1602
Db 3439 TATTGGAGGCGACTTTCAGCAGTTCCAGAACGATGATGAATTAAGCCGGCAGGCAAC 3498
Qy 1603 GGAATTGAAAACCTAAAGAACCTGTATCATGACTACTCTTGAGACTGTACGAATATTTCT 1662
Db 3499 GGAATTGAAAACCTAAAGAACCTGTATCATGACTACTCTTGAGACTGTACGAATATTTCT 3558
Qy 1663 GACAGAGCAGCCTTTTGGAAAGGACTAGAGAACTCTACCAGGAGCCAGAGAGCTGCCTCC 1722
Db 3559 GACAGAGCAGCCTTTTGGAAAGGACTAGAGAACTCTACCAGGAGCCAGAGAGCTGCCTCC 3618
Qy 1723 TGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATAC 1782
Db 3619 TGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATAC 3678
Qy 1783 TGAGTGGGAAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCT 1842
Db 3679 TGAGTGGGAAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCT 3738
Qy 1843 TGAAAGACTCCAGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGC 1902
Db 3739 TGAAAGACTCCAGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGC 3798
Qy 1903 TGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAGA 1962
Db 3799 TGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAGA 3858
Qy 1963 TCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAAGAGAACGTGAG 2022
Db 3859 TCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAAGAGAACGTGAG 3918
Qy 2023 CCACGTCAATGACCTTGTCTGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAA 2082
Db 3919 CCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAA 3978
Qy 2083 CCTCAGCACTCTGGAAGAC 2101
Db 3979 CCTCAGCACTCTGGAAGAC 3997

QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCACTGAC 1020
Db 1851 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCACTGAC 1910
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
Db 1911 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1970
QY 1081 ACTTGAAAAGAGTACAGCACAG----- 1102
Db 1971 ACTTGAAAAGAGTACAGCACAGGAAACTGAAATAGCAGTTCAAGCTAAACAACCGGATGT 2030
QY 1103 ----- 1102
Db 2031 GGAAGAGATTTGTCTAAAGGCGCAGCATCTGAGCTCTGAGTGTGAAGGCGGTAAACCGTTTACTTCAAGA 2150
QY 1103 ----- 1102
Db 2091 GAAGAGGAGTTAGAAGATCTGAGCTCTGAGTGTGAAGGCGGTAAACCGTTTACTTCAAGA 2150
QY 1103 ----- 1102
Db 2151 GCTGAGGGCAAGCAGCCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTCTCCTAC 2210
QY 1103 ----- 1102
Db 2211 TCAGACTGTACTCTGGTGACACAACCTGTGTTACTAAGGAAACTGCCATCTCCAAACT 2270
QY 1103 ----- 1102
Db 2271 AGAAATGCCATCTTCTTGATGTTGGAGGTACCTGCTCTGGCAGATTTCACCGGGCTTG 2330
QY 1103 ----- 1102
Db 2331 GACAGAACTTACCGACTGGCTTTCTCTGCTTGATCAAGTTATAAAATCACAGAGGGTGAT 2390
QY 1103 ----- 1102
Db 2391 GGTGGTGACCTTGAGGATATCAACGAGATGATCATCAAGCAGAGGCAACAATGCAGGA 2450
QY 1103 ----- 1102
Db 2451 TTGTGAAACAGAGGCGTCCCAGTTGGAAAGAACTCATTACCGCTGCCCAAATTTGAAAAA 2510
QY 1103 ----- 1102
Db 2511 CAAGACCAGCAATCAAGAGGCTAGAAACAATCATACGGATCGAATTGAAAGAATTCAGAA 2570
QY 1103 ----- 1102
Db 2571 TCAGTGGGATGAAGTACAAGAACACCTTCAGAACCGGAGGCAACAGTTGAATGAAATGTT 2630
QY 1103 ----- 1102
Db 2631 AAAGGATTCAAACAATGGCTGGAAGCTAAGGAAGAAGCTGAGCAGGTCTTAGGACAGGC 2690
QY 1103 ----- 1102
Db 2691 CAGAGCCAAGCTTGAGTCAATGGAAGGAGGCTCCCTATACAGTAGATGCAATCCAAAAAGAA 2750
QY 1103 ----- 1102
Db 2751 AATCACAGAAACCAAGCAGTTGGCCAAAGACCTCCGCCAGTGGCAGACAAATGTAGATGT 2810
QY 1103 ----- 1102
Db 2811 GGCAATGACTTGGCCCTGAAACTTCTCCGGGATTAATCTGCAGATGATACCAGAAAAAGT 2870
QY 1103 ----- 1102
Db 2871 CCACATGATAACAGAGAAATATCAATGCCTCTTGGAGAAGCATTTCATAAAAGGGTGATGA 2930

QY 1103 -----ACTCATAGATTACTGCAACAGTTCCCCCTGGACCTTGA 1140
Db 2931 GCGAGAGGCTGCTTTGGAAGAAACTCATAGATTACTGCAACAGTTCCCCCTGGACCTTGA 2990
QY 1141 AAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTAC 1200
Db 2991 AAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTAC 3050
QY 1201 CCGTAAGGAAAGGCTCCTAGAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 1260
Db 3051 CCGTAAGGAAAGGCTCCTAGAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 3110
QY 1261 AGACCTCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAACAACTGGATGAAAAACAG 1320
Db 3111 AGACCTCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAACAACTGGATGAAAAACAG 3170
QY 1321 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAAGTCTCTCAACATTAGGTCCCA 1380
Db 3171 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAAGTCTCTCAACATTAGGTCCCA 3230
QY 1381 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCCA 1440
Db 3231 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCCA 3290
QY 1441 TTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGCT 1500
Db 3291 TTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGCT 3350
QY 1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCGGCGGAGGCACTTATGGAGGCGACTTTCC 1560
Db 3351 GTGGCTACAGCTGAAAGATGATGAATTAAGCGGCGGAGGCACTTATGGAGGCGACTTTCC 3410
QY 1561 AGCAGTTCAAGAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTTGAAAACTAAAGA 1620
Db 3411 AGCAGTTCAAGAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTTGAAAACTAAAGA 3470
QY 1621 ACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGA 1680
Db 3471 ACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGA 3530
QY 1681 AGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCGAGAA 1740
Db 3531 AGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCGAGAA 3590
QY 1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAAATTGAA 1800
Db 3591 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAAATTGAA 3650
QY 1801 CCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 1860
Db 3651 CCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 3710
QY 1861 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATC 1920
Db 3711 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATC 3770
QY 1921 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAA 1980
Db 3771 CTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAA 3830
QY 1981 GGCACCTTCGAGGAGAAAATTGGCGCTCTGAAAGAGAACGCTGAGCCACGTCATGACCTTGC 2040
Db 3831 GGCACCTTCGAGGAGAAAATTGGCGCTCTGAAAGAGAACGCTGAGCCACGTCATGACCTTGC 3890
QY 2041 TCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAGCTCTTGAAGA 2100
Db 3891 TCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAGCTCTTGAAGA 3950
QY 2101 C 2101
Db 3951 C 3951

RESULT 12
AAD37260
ID AAD37260 standard; DNA; 4414 BP.
XX AC AAD37260;
XX DT 21-AUG-2002 (first entry)
XX Adeno-associated virus vector plasmid, AAV-MCK-3447.
DE Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX Chimeric - Homo sapiens.
OS Chimeric - Unidentified.
XX PN WO200183695-A2.
XX PD 08-NOV-2001.
XX PF 27-APR-2001; 2001WO-US13677.
XX PR 28-APR-2000; 2000US-200777P.
XX PA (XIAO/) XIAO X.
XX PI Xiao X;
XX WPI; 2002-049342/06.
XX PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX Example 1; Page 65-66; 71pp; English.
XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence.
XX Sequence 4414 BP; 1255 A; 1075 C; 1086 G; 998 T; 0 other;
SQ Query Match 53.0%; Score 1114; DB 24; Length 4414;
Best Local Similarity 97.0%; Pred. No. 7.5e-301;
Matches 1135; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTACACCACCTCTGACCCCTACACGGAGCCCAT 60
Db 1657 GAGCTATGCCTACACACAGGCTGCTTATGTACACCACCTCTGACCCCTACACGGAGCCCAT 1716
QY 61 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 120
Db 1717 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 1776
QY 121 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
Db 1777 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1836
QY 181 TGCTGAGGACACATTGCAACGACCAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 240
Db 1837 TGCTGAGGACACATTGCAACGACCAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 1896

QY 241 CCAGTTTTCATACTCATFAGGGGTACATGATGATTTGACAGCCCATCAGGCGCGGTGG 300
Db 1897 CCAGTTTTCATACTCATFAGGGGTACATGATGATTTGACAGCCCATCAGGCGCGGTGG 1956
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGTAAGCAAGGAAATTTATCAGAAGATGAAGA 360
Db 1957 TAATATTCTACAATTGGGAAGTAAGCTGATTGTAAGCAAGGAAATTTATCAGAAGATGAAGA 2016
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATCAAGATGGGAATGCCTCAGGGTAGC 420
Db 2017 AACTGAAGTACAAGAGCAGATGAATCTCTAAATCAAGATGGGAATGCCTCAGGGTAGC 2076
QY 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 480
Db 2077 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 2136
QY 481 GAAAGAGTTGAATGACTGCTGCTAAACAAACAGAAAGAAACAAGGAAATGGAGGAAGA 540
Db 2137 GAAAGAGTTGAATGACTGCTGCTAAACAAACAGAAAGAAACAAGGAAATGGAGGAAGA 2196
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAACAACATAAGGTGCTTCA 600
Db 2197 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAACAACATAAGGTGCTTCA 2256
QY 601 AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGGTAGT 660
Db 2257 AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGGTAGT 2316
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAAACAACCTTAAGGTATTGGG 720
Db 2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAAACAACCTTAAGGTATTGGG 2376
QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGCCGCTGGGTCTCTTTTACAAGACAT 780
Db 2377 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGCCGCTGGGTCTCTTTTACAAGACAT 2436
QY 781 CCTTCTCAAATGGCAACGCTTTACTGAAGAACAGTGCGCTTTTGTGATGCGATGGCTTTTCA 840
Db 2437 CCTTCTCAAATGGCAACGCTTTACTGAAGAACAGTGCGCTTTTGTGATGCGATGGCTTTTCA 2496
QY 841 AAAAGAAGATGCAGTGAACAAGATTCAACAACCTGGCTTTAAAGATCAAAATGAAATGTT 900
Db 2497 AAAAGAAGATGCAGTGAACAAGATTCAACAACCTGGCTTTAAAGATCAAAATGAAATGTT 2556
QY 901 ATCAAGTCTTCAAAAACCTGGCGCTTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT 960
Db 2557 ATCAAGTCTTCAAAAACCTGGCGCTTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT 2616
QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db 2617 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2676
QY 1021 CCAGAAGACCGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1080
Db 2677 CCAGAAGACCGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAA 2736
QY 1081 ACTTGAAGAAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCTCCCTGGACCTGGA 1140
Db 2737 ACTTGAAGAAGTACAGCACAGACCCCTTGAAGAAGACTCCAGGHAATCAAGAGGCCACCGA 2796
QY 1141 AAAGTTTCTTGGCTGGCTTACAGAAGCTGA 1170
Db 2797 TGAGCTGGACCTCAAGCTGCGCCCAAGCTGA 2826
RESULT 13
ABK81997
ID ABK81997 standard; DNA; 5417 BP.
XX AC ABK81997;
XX DT 13-AUG-2002 (first entry)
XX

DE DNA encoding mini-dystrophin protein deltaR4-R23.
XX
KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200229056-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US311126.
XX
PR 06-OCT-2000; 2000US-238848P.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Chamberlain JS, Harper SQ;
XX
DR WPI; 2002-435334/46.
XX
XX
PT A composition for preparing therapeutic drugs, has a mini-dystrophin
PT peptide comprising a specific number of spectrin-like repeat domains,
PT or a nucleic acid sequence encoding the mini-dystrophin peptide -
XX
XX Disclosure; Fig 12; 145pp; English.
XX
XX The invention describes a composition comprising a mini-dystrophin
CC peptide comprising a spectrin-like repeat domain, where the domain
CC comprises n spectrin-like repeats, and contains no more than n
CC spectrin-like repeats, where n is an even number between 4-24, or a
CC nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin
CC peptide or the polynucleotide encoding it is useful as a medicament,
CC for preparing a drug for therapeutic application and in the preparation
CC of a composition for treatment of muscle disease, e.g. Duchenne's
CC muscular dystrophy (DMD). This sequence represents a mini-dystrophin
CC sequence of the invention.
XX
SQ Sequence 5417 BP; 1700 A; 1192 C; 1182 G; 1343 T; 0 other;

Query Match 52.5%; Score 1103.4; DB 24; Length 5417;
Best Local Similarity 99.9%; Pred. No. 7.8e-298;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGCTATGCTTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCAT 60
DB 1099 GAGCTATGCTTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCAT 1158

QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 120
DB 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 1218

QY 121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTCGTGGCTTCTTC 180
DB 1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTCGTGGCTTCTTC 1278

QY 181 TGCTGAGGACACATTGCAAGCACAAGCAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
DB 1279 TGCTGAGGACACATTGCAAGCACAAGCAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338

QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGGATTGTGACAGCCCATCAGGGCCGGTTGG 300
DB 1339 CCAGTTTCATCTCATGAGGGGTACATGATGGATTGTGACAGCCCATCAGGGCCGGTTGG 1398

QY 301 TAATATTCTACATTTGGGAAGTAAGCTGATTGGAACAGGAAATATCAGAAGATGAAGA 360
DB 1399 TAATATTCTACATTTGGGAAGTAAGCTGATTGGAACAGGAAATATCAGAAGATGAAGA 1458

QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
DB 1459 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1518

QY 421 TAGCATGAAAAACAAAGCAATTACATAGAGTTTTTAATGGATCTCCAGAACTCAGAAACT 480
DB 1519 TAGCATGAAAAACAAAGCAATTACATAGAGTTTTTAATGGATCTCCAGAACTCAGAAACT 1578
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAAACAAGGAAATGGAGGAAGA 540
DB 1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAAACAAGGAAATGGAGGAAGA 1638
QY 541 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAACAACATAAGGTGCTTCA 600
DB 1639 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAACAACATAAGGTGCTTCA 1698
QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGTAGT 660
DB 1699 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGTAGT 1758
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAGAACAACTTAAGGTATTGGG 720
DB 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAGAACAACTTAAGGTATTGGG 1818
QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTCTTTTACAAGACAT 780
DB 1819 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTCTTTTACAAGACAT 1878
QY 781 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 840
DB 1879 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 1938
QY 841 AAAAGAGATGCAGTGAACAAGATTCAACAACACTGGCTTTTAAAGATCAAAATGAAATGTT 900
DB 1939 AAAAGAGATGCAGTGAACAAGATTCAACAACACTGGCTTTTAAAGATCAAAATGAAATGTT 1998
QY 901 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT 960
DB 1999 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT 2058
QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
DB 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2118
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTAGTCCAAA 1080
DB 2119 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTAGTCCAAA 2178
QY 1081 ACTTGAAAGAGTACAGCACAGACT 1105
DB 2179 ACTTGAAAGAGTACAGCACAGATT 2203

RESULT 14
ABK82005
ID ABK82005 standard; DNA; 11241 BP.
XX
AC ABK82005;
XX
DT 13-AUG-2002 (first entry)
XX
DE cDNA encoding human dystrophin, full length HDMD.
XX
KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; human; ds.
XX
OS Homo sapiens.
OS
PN WO200229056-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US311126.
XX
PR 06-OCT-2000; 2000US-238848P.
XX
PA (UNMI) UNIV MICHIGAN.

Chamberlain JS, Harper SQ;
WPI; 2002-435334/46.

A composition for preparing therapeutic drugs, has a mini-dystrophin peptide comprising a specific number of spectrin-like repeat domains, or a nucleic acid sequence encoding the mini-dystrophin peptide -

Example 2; Fig 23; 145pp; English.

The invention describes a composition comprising a mini-dystrophin peptide comprising a spectrin-like repeat domain, where the domain comprises n spectrin-like repeats, and contains no more than n spectrin-like repeats, where n is an even number between 4-24, or a nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the polynucleotide encoding it is useful as a medicament, for preparing a drug for therapeutic application and in the preparation of a composition for treatment of muscle disease, e.g. Duchenne's muscular dystrophy (DMD). This sequence represents a human dystrophin polynucleotide sequence used in the creation of the mini-dystrophin peptides of the invention.

Sequence 11241 BP; 3738 A; 2325 C; 2656 G; 2522 T; 0 other;

Query Match 52.5%; Score 1103.4; DB 24; Length 11241;
Best Local Similarity 99.9%; Pred. No. 1.2e-297;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

| | | | |
|----|------|--|------|
| QY | 1 | GAGCTATGCCTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCAT | 60 |
| Db | 1099 | GAGCTATGCCTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCAT | 1158 |
| QY | 61 | TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG | 120 |
| Db | 1159 | TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG | 1218 |
| QY | 121 | TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTC | 180 |
| Db | 1219 | TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTC | 1278 |
| QY | 181 | TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA | 240 |
| Db | 1279 | TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA | 1338 |
| QY | 241 | CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGSCGGTTGG | 300 |
| Db | 1339 | CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGSCGGTTGG | 1398 |
| QY | 301 | TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTTATCAGAAGATGAAGA | 360 |
| Db | 1399 | TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTTATCAGAAGATGAAGA | 1458 |
| QY | 361 | AACTGAAGTACAAGACGACATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC | 420 |
| Db | 1459 | AACTGAAGTACAAGACGACATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC | 1518 |
| QY | 421 | TAGCATGGAATAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAAAC | 480 |
| Db | 1519 | TAGCATGGAATAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAAAC | 1578 |
| QY | 481 | GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAGGAAAATGGAGGAAGA | 540 |
| Db | 1579 | GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAGGAAAATGGAGGAAGA | 1638 |
| QY | 541 | GCCTCTGGACCTGATCTTTGAAGACCTAAAACGCCCAAGTACAACAATAGGTGCTTCA | 600 |
| Db | 1639 | GCCTCTGGACCTGATCTTTGAAGACCTAAAACGCCCAAGTACAACAATAGGTGCTTCA | 1698 |
| QY | 601 | AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGTAGT | 660 |
| Db | 1699 | AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGTAGT | 1758 |

| | | | |
|-----------|---|---|------|
| QY | 661 | TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAACAACCTTAAGGTATTGGG | 720 |
| DB | 1759 | TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAACAACCTTAAGGTATTGGG | 1818 |
| QY | 721 | AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTCTTTTACAAGACAT | 780 |
| DB | 1819 | AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTCTTTTACAAGACAT | 1878 |
| QY | 781 | CCTTCTCAAATGGCAACGTCCTTACTGAAGAACAGTGCCCTTTTAGTGATGGCTTTTCAGA | 840 |
| DB | 1879 | CCTTCTCAAATGGCAACGTCCTTACTGAAGAACAGTGCCCTTTTAGTGATGGCTTTTCAGA | 1938 |
| QY | 841 | AAAAGAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT | 900 |
| DB | 1939 | AAAAGAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT | 1998 |
| QY | 901 | ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT | 960 |
| DB | 1999 | ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT | 2058 |
| QY | 961 | GGGCCAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAAGTCAGTGAC | 1020 |
| DB | 2059 | GGGCCAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAAGTCAGTGAC | 2118 |
| QY | 1021 | CCAGAAGACCGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA | 1080 |
| DB | 2119 | CCAGAAGACCGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA | 2178 |
| QY | 1081 | ACTTGAAAAGAGTACAGCACAGACT | 1105 |
| DB | 2179 | ACTTGAAAAGAGTACAGCACAGATT | 2203 |
| | | | |
| RESULT 15 | | | |
| ABK82002 | | | |
| ID | ABK82002 standard; DNA; 11443 BP. | | |
| XX | | | |
| AC | ABK82002; | | |
| XX | | | |
| DT | 13-AUG-2002 (first entry) | | |
| XX | | | |
| DE | DNA encoding mini-dystrophin protein deltaR9-R16. | | |
| XX | | | |
| KW | Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease; | | |
| KW | Duchenne's muscular dystrophy; DMD; dystrophin; ds. | | |
| XX | | | |
| OS | Homo sapiens. | | |
| OS | Synthetic. | | |
| XX | | | |
| PN | WO200229056-A2. | | |
| PD | | | |
| XX | 11-APR-2002. | | |
| XX | | | |
| Pf | 04-OCT-2001; 2001WO-US31126. | | |
| XX | | | |
| PR | 06-OCT-2000; 2000US-238848P. | | |
| XX | | | |
| PA | (UNMI) UNIV MICHIGAN. | | |
| XX | | | |
| PI | Chamberlain JS, Harper SQ; | | |
| XX | | | |
| DR | WPI; 2002-435334/46. | | |
| XX | | | |
| PT | A composition for preparing therapeutic drugs, has a mini-dystrophin | | |
| PT | peptide comprising a specific number of spectrin-like repeat domains, | | |
| PT | or a nucleic acid sequence encoding the mini-dystrophin peptide - | | |
| XX | | | |
| PS | Disclosure; Fig 17; 145pp; English. | | |
| XX | | | |
| CC | The invention describes a composition comprising a mini-dystrophin | | |
| CC | peptide comprising a spectrin-like repeat domain, where the domain | | |
| CC | comprises n spectrin-like repeats, and contains no more than n | | |
| CC | spectrin-like repeats, where n is an even number between 4-24, or a | | |

CC nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin
CC peptide or the polynucleotide encoding it is useful as a medicament,
CC for preparing a drug for therapeutic application and in the preparation
CC of a composition for treatment of muscle disease, e.g. Duchenne's
CC muscular dystrophy (DMD). This sequence represents a mini-dystrophin
CC sequence of the invention.
XX
SQ Sequence 11443 BP; 3707 A; 2339 C; 2502 G; 2895 T; 0 other;

Query Match 52.5%; Score 1103.4; DB 24; Length 11443;
Best Local Similarity 99.9%; Pred. No. 1.2e-297;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGCTATGCCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCCTACACGGAGCCCAATT 60
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1099 GAGCTATGCCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCCTACACGGAGCCCAATT 1158
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 61 TCCTTCACAGCAATTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG 120
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1159 TCCTTCACAGCAATTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG 1218
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCCTTC 180
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCCTTC 1278
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 181 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1279 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 241 CCAGTTTTCATACATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTGG 300
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1339 CCAGTTTTCATACATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTGG 1398
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAATAATTATCAGAAGATGAAGA 360
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAATAATTATCAGAAGATGAAGA 1458
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1519 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 1578
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAACACAGGAAAAATGGAGGAAGA 540
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAACACAGGAAAAATGGAGGAAGA 1638
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 541 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAAACACATAAGGTGCTTCA 600
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1639 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAAACACATAAGGTGCTTCA 1698
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACTGGTGGTAGT 660
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1699 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACTGGTGGTAGT 1758
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1818
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 780
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1819 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 1878
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 781 CCTTCTCAAATGGCAACGTCCTTACTGAAGAACAGTGCCTTTTATGTCATGGCTTTCAGA 840
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1879 CCTTCTCAAATGGCAACGTCCTTACTGAAGAACAGTGCCTTTTATGTCATGGCTTTCAGA 1938
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 841 AAAAGAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTAAAGATCAAAATGAATGTT 900
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1939 AAAAGAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTAAAGATCAAAATGAATGTT 1998
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 901 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT 960
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1999 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT 2058
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 961 GGGCAAACTGTATTCTACTCAAAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGAC 1020
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2059 GGGCAAACTGTATTCTACTCAAAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGAC 2118
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2119 CCAGAAGACGGAAGCATGGCTGGATAACTTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 2178
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 1081 ACTTGAAGAAGAGTACAGCACAGACT 1105
DB ||||||||||||||||||||||||||||||||
2179 ACTTGAAGAAGAGTACAGCACAGATT 2203
DB ||||||||||||||||||||||||||||||||

Search completed: February 1, 2004, 12:03:57
Job time : 536.045 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OMnucleic - nucleic search, using sw model
Run on: February 1, 2004, 10:14:55 ; Search time 7484.44 Seconds
(without alignments)
11483.979 Million cell updates/sec

Title: US-09-845-416-2_COPY_900_3000
Perfect score: 2101
Sequence: 1 gagctatgcctacacacagg.....acctcagcactctggaagac 2101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID | Description |
|------------|--------|-------|--------|----|----------|---------------------|
| 1 | 1293 | 61.5 | 8689 | 6 | AX538622 | AX538622 Sequence |
| 2 | 1131 | 53.8 | 5952 | 6 | AR304538 | AR304538 Sequence |
| 3 | 1131 | 53.8 | 5952 | 6 | AX114289 | AX114289 Sequence |
| 4 | 1103.4 | 52.5 | 5417 | 6 | AX538619 | AX538619 Sequence |
| 5 | 1103.4 | 52.5 | 11443 | 6 | AX538624 | AX538624 Sequence |
| 6 | 1103.4 | 52.5 | 12057 | 6 | AX538627 | AX538627 Sequence |
| 7 | 1103.4 | 52.5 | 13957 | 6 | AX409637 | AX409637 Sequence |
| 8 | 1103.4 | 52.5 | 13957 | 6 | AX538581 | AX538581 Sequence |
| 9 | 1103.4 | 52.5 | 13957 | 9 | HUMDYS | M18533 Homo sapien |
| 10 | 1103.4 | 52.5 | 13977 | 6 | AR220819 | AR220819 Sequence |
| 11 | 1100.2 | 52.4 | 12446 | 9 | HSDMDR | X14298 Human mRNA |
| 12 | 997 | 47.5 | 5339 | 6 | AX538620 | AX538620 Sequence |
| 13 | 996 | 47.4 | 5462 | 6 | AX538621 | AX538621 Sequence |
| 14 | 929 | 44.2 | 13887 | 4 | AF070485 | AF070485 Canis fam |
| 15 | 911.6 | 43.4 | 4402 | 6 | E30220 | E30220 Shortened d |
| 16 | 866.6 | 41.2 | 3275 | 10 | MUSDYS | M18025 Mouse dyster |
| 17 | 866.6 | 41.2 | 13815 | 6 | AX306153 | AX306153 Sequence |
| 18 | 866.6 | 41.2 | 13815 | 6 | AX538582 | AX538582 Sequence |
| 19 | 866.6 | 41.2 | 13815 | 10 | MUSDYSA | M68859 Mouse dyster |
| 20 | 866.6 | 41.2 | 19307 | 6 | AR093392 | AR093392 Sequence |
| 21 | 866.6 | 41.2 | 19307 | 6 | AR142592 | AR142592 Sequence |
| 22 | 718.8 | 34.2 | 4075 | 6 | E30221 | E30221 Shortened d |
| 23 | 714.6 | 34.0 | 4402 | 6 | E30219 | E30219 Shortened d |
| 24 | 635.8 | 30.3 | 13575 | 5 | GGDYS | X13369 Chicken mRN |
| 25 | 624.2 | 29.7 | 630 | 9 | HSDMDF1 | X06179 Human fetal |
| 26 | 527.8 | 25.1 | 1966 | 9 | HSDMDA1 | X06178 Human adult |
| 27 | 409.2 | 19.5 | 3747 | 6 | E30218 | E30218 Shortened d |
| 28 | 387 | 18.4 | 387 | 6 | AX538612 | AX538612 Sequence |
| 29 | 361.4 | 17.2 | 10302 | 6 | AX538583 | AX538583 Sequence |
| 30 | 361.4 | 17.2 | 10302 | 9 | HSMUPS | X69086 H.sapiens m |
| 31 | 350.4 | 16.7 | 10705 | 10 | RNAJ2967 | AJ002967 Rattus no |
| 32 | 348 | 16.6 | 348 | 6 | AX538611 | AX538611 Sequence |
| 33 | 344.6 | 16.4 | 1898 | 10 | BC024140 | BC024140 Mus muscu |
| 34 | 344.4 | 16.4 | 3521 | 5 | AF339031 | AF339031 Danio rer |
| 35 | 343 | 16.3 | 2394 | 10 | BC050183 | BC050183 Mus muscu |
| 36 | 343 | 16.3 | 6045 | 6 | A63605 | A63605 Sequence 7 |
| 37 | 343 | 16.3 | 6045 | 6 | AR281528 | AR281528 Sequence |
| 38 | 343 | 16.3 | 10320 | 6 | A63607 | A63607 Sequence 9 |
| 39 | 343 | 16.3 | 10320 | 6 | AR281529 | AR281529 Sequence |
| 40 | 336.6 | 16.0 | 6059 | 6 | AX107972 | AX107972 Sequence |
| 41 | 336.6 | 16.0 | 11096 | 6 | AX538584 | AX538584 Sequence |
| 42 | 336.6 | 16.0 | 11096 | 10 | MMY12229 | Y12229 M.musculus |
| 43 | 333 | 15.8 | 333 | 6 | AX538589 | AX538589 Sequence |
| 44 | 327 | 15.6 | 327 | 6 | AX538588 | AX538588 Sequence |
| 45 | 322.4 | 15.3 | 333 | 6 | AX538590 | AX538590 Sequence |

ALIGNMENTS

RESULT 1
AX538622
LOCUS AX538622 8689 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 42 from Patent WO0229056.
ACCESSION AX538622
VERSION AX538622.1 GI:25271171
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.

REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 42 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

| FEATURES | | Location/Qualifiers | |
|---|-----------------------------|---|------|
| source | 1..8689 | /organism="synthetic construct" | |
| | | /mol_type="genomic DNA" | |
| | | /db_xref="taxon:32630" | |
| | | /note="Synthetic" | |
| SE COUNT | 2721 a 1804 c 1861 g 2303 t | | |
| IGIN | | | |
| Query Match 61.5%; Score 1293; DB 6; Length 8689; | | | |
| Best Local Similarity 72.5%; Pred. No. 3.3e-296; | | | |
| Matches 2101; Conservative 0; Mismatches 0; Indels 798; Gaps 1; | | | |
| QY | 1 | GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT | 60 |
| Db | 1099 | GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT | 1158 |
| QY | 61 | TCCTTCACAGCAATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGTAGGAGAG | 120 |
| Db | 1159 | TCCTTCACAGCAATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGTAGGAGAG | 1218 |
| QY | 121 | TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAATATATCGTGGCTTCTTTC | 180 |
| Db | 1219 | TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAATATATCGTGGCTTCTTTC | 1278 |
| QY | 181 | TGCTGAGGACACATTTGCAAGCACAAAGGACAGATTTCTAATGATGTGGAAGTGTGAAAGA | 240 |
| Db | 1279 | TGCTGAGGACACATTTGCAAGCACAAAGGACAGATTTCTAATGATGTGGAAGTGTGAAAGA | 1338 |
| QY | 241 | CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG | 300 |
| Db | 1339 | CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG | 1398 |
| QY | 301 | TAATATTCTACAATTTGGGAAGTAAGCTGATTTGGAACAGGAAATTTATCAGAAGATGAAGA | 360 |
| Db | 1399 | TAATATTCTACAATTTGGGAAGTAAGCTGATTTGGAACAGGAAATTTATCAGAAGATGAAGA | 1458 |
| QY | 361 | AACTGAAGTACAAGACAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC | 420 |
| Db | 1459 | AACTGAAGTACAAGACAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC | 1518 |
| QY | 421 | TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT | 480 |
| Db | 1519 | TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT | 1578 |
| QY | 481 | GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGGAAATGGAGGAAGA | 540 |
| Db | 1579 | GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGGAAATGGAGGAAGA | 1638 |
| QY | 541 | GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAACAACATAAGGTGCTTCA | 600 |
| Db | 1639 | GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAACAACATAAGGTGCTTCA | 1698 |
| QY | 601 | AGAAGATCTAGAACCAAGAACAGTCAAGGTCAATTTCTCTCACTCACATGGTGGTGTAGT | 660 |
| Db | 1699 | AGAAGATCTAGAACCAAGAACAGTCAAGGTCAATTTCTCACTCACATGGTGGTGTAGT | 1758 |
| QY | 661 | TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG | 720 |
| Db | 1759 | TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG | 1818 |
| QY | 721 | AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT | 780 |
| Db | 1819 | AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT | 1878 |
| QY | 781 | CCTTCTCAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA | 840 |
| Db | 1879 | CCTTCTCAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA | 1938 |
| QY | 841 | AAAAGAGATGCAGTGAACAAGATTCACACAACCTGGCTTTAAAGATCAAAATGAAATGTT | 900 |
| Db | 1939 | AAAAGAGATGCAGTGAACAAGATTCACACAACCTGGCTTTAAAGATCAAAATGAAATGTT | 1998 |
| QY | 901 | ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCCGGATCTAGAAAAAGAAAAAGCAATCCAT | 960 |
| Db | 1999 | ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCCGGATCTAGAAAAAGAAAAAGCAATCCAT | 2058 |
| QY | 961 | GGGCAAACTGTATTCTACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC | 1020 |
| Db | 2059 | GGGCAAACTGTATTCTACTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC | 2118 |
| QY | 1021 | CCAGAACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAAA | 1080 |
| Db | 2119 | CCAGAACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAAA | 2178 |
| QY | 1081 | ACTTGAAAAGAGTACAGCACAG | 1102 |
| Db | 2179 | ACTTGAAAAGAGTACAGCACAGATTTCAACAGACCTGACCTAGCTCCTGGACTGACCCAC | 2238 |
| QY | 1103 | | 1102 |
| Db | 2239 | TATTGGAGCCTCTCCTACTCAGACTGTACTCTGGTGACACAACCTGTGGTTACTAAGGA | 2298 |
| QY | 1103 | | 1102 |
| Db | 2299 | AACTGCCATCTCCAAAACCTAGAAATGCCATCTTCTTGTATGTTGGAGGTACCTGCTCTGGC | 2358 |
| QY | 1103 | | 1102 |
| Db | 2359 | AGATTTCAACCGGGCTTGGACAGAACTTACCGACTGGCTTCTCTGTGATCAAGTTAT | 2418 |
| QY | 1103 | | 1102 |
| Db | 2419 | AAAATCACAGAGGGTGATGGTGGTGACCTTGAGGATATCAACGAGATGATCATCAAGCA | 2478 |
| QY | 1103 | | 1102 |
| Db | 2479 | GAAGGCAACAATGCAGGATTTTGGAAACAGAGGCGTCCCAGTTTGGAGAACTCATTACCGC | 2538 |
| QY | 1103 | | 1102 |
| Db | 2539 | TGCCCAAAATTTGAAAAACAAGACCAGCAATCAAGAGGCTAGAACAACTATTACGGATCG | 2598 |
| QY | 1103 | | 1102 |
| Db | 2599 | AATTGAAAAGATTCAGAATCAGTGGGATGAAGTACAAGAACACCTTCAGAACCCGAGGCA | 2658 |
| QY | 1103 | | 1102 |
| Db | 2659 | ACAGTTGAATGAATGTTAAAGGATTCACACAATGGCTGGAAGCTAAGGAAGAAGCTGA | 2718 |
| QY | 1103 | | 1102 |
| Db | 2719 | GCAGGTCTTAGGACAGGCCAGCCAAAGCTTGAGTCATGGAAGGAGGCTCCCTATACAGT | 2778 |
| QY | 1103 | | 1102 |
| Db | 2779 | AGATGCAATCCAAAAGAAAAATCACAGAAACCAAGCAGTTGGCCAAAGACCTCCGCCAGTG | 2838 |
| QY | 1103 | | 1102 |
| Db | 2839 | GCAGACAAATGTAGATGTGGCAATGACTTGGCCCTGAAACTTCTCCGGGATTATTCTGTC | 2898 |
| QY | 1103 | | 1102 |
| Db | 2899 | AGATGATACCAGAAAAGTCCACATGATAACAGAGAAATATCAATGCCTCTTGGAGAAGCAT | 2958 |
| QY | 1103 | -----ACTCATAGATTACTGCAACA | 1122 |
| Db | 2959 | TCATAAAAAGGTTGAGTGAGCGCAGAGGCTGCTTTTGGAAAGAACTCATAGATTACTGCAACA | 3018 |
| QY | 1123 | GTTCCCTTGGACCTGGAAAAAGTTTCTTGGCTGGCTTACAGAAGCTGAAACAACTGCCAA | 1182 |
| Db | 3019 | GTTCCCTTGGACCTGGAAAAAGTTTCTTGGCTGGCTTACAGAAGCTGAAACAACTGCCAA | 3078 |
| QY | 1183 | TGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCTAGAAAGCTCCAGGGAGTAAAAAGA | 1242 |

Db 3079 TGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAGACTCCAAGGGAGTAAAGA 3138
QY 1243 GCTGATGAAACAATGGCAAGACCTCCAAAGTGAAATTAAGCTCACACAGATGTTTATCA 1302
Db 3139 GCTGATGAAACAATGGCAAGACCTCCAAAGTGAAATTAAGCTCACACAGATGTTTATCA 3198
QY 1303 CAACCTGGATGAAACAGCCAAAAATCCTGAGATCCCTGGAAAGTCCGATGATGCAGT 1362
Db 3199 CAACCTGGATGAAACAGCCAAAAATCCTGAGATCCCTGGAAAGTCCGATGATGCAGT 3258
QY 1363 CCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTC 1422
Db 3259 CCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTC 3318
QY 1423 TCTCAACATTAGGTCCCATTTGGAAGCCAGTCTGACCAAGTGGAAGCGTCTGCACCTTTC 1482
Db 3319 TCTCAACATTAGGTCCCATTTGGAAGCCAGTCTGACCAAGTGGAAGCGTCTGCACCTTTC 3378
QY 1483 TCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACC 1542
Db 3379 TCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACC 3438
QY 1543 TATTGGAGGCGACTTCCAGCAGTTTCAGAGCAGTTCAGAAAGCAGATGTACATAGGGCCTTCAAGAG 1602
Db 3439 TATTGGAGGCGACTTCCAGCAGTTTCAGAGCAGTTCAGAAAGCAGATGTACATAGGGCCTTCAAGAG 3498
QY 1603 GGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCT 1662
Db 3499 GGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCT 3558
QY 1663 GACAGAGCAGCCTTTTGAAGGACTAGAGAACTCTACCAGGAGCCCAAGAGAGCTGCCTCC 1722
Db 3559 GACAGAGCAGCCTTTTGAAGGACTAGAGAACTCTACCAGGAGCCCAAGAGAGCTGCCTCC 3618
QY 1723 TGAGGAGAGAGCCCAAGAAATGTCACTCGGCTTACGAAAGCAGGCTGAGGAGGTCAATAC 1782
Db 3619 TGAGGAGAGAGCCCAAGAAATGTCACTCGGCTTACGAAAGCAGGCTGAGGAGGTCAATAC 3678
QY 1783 TGAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCT 1842
Db 3679 TGAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCT 3738
QY 1843 TGAAGAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAAAGC 1902
Db 3739 TGAAGAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAAAGC 3798
QY 1903 TGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAGA 1962
Db 3799 TGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAGA 3858
QY 1963 TCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAGAGAAACGTGAG 2022
Db 3859 TCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAGAGAAACGTGAG 3918
QY 2023 CCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTGGGCATTCAGCTCTCACCCGTATAA 2082
Db 3919 CCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTGGGCATTCAGCTCTCACCCGTATAA 3978
QY 2083 CCTCAGCACTCTGGAAGAC 2101
Db 3979 CCTCAGCACTCTGGAAGAC 3997

RESULT 2
AR304538
LOCUS AR304538 5952 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 1 from patent US 6544786.
ACCESSION AR304538
VERSION AR304538.1 GI:31693691
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

Unclassified.
REFERENCE 1 (bases 1 to 5952)
AUTHORS Xiao,X. and Liu,P.X.
TITLE Method and vector for producing and transferring trans-spliced peptides
JOURNAL Patent: US 6544786-A 1 08-APR-2003;
FEATURES Location/Qualifiers
source 1. .5952
BASE COUNT 1860 a 1344 c 1410 g 1338 t
ORIGIN
Query Match 53.8%; Score 1131; DB 6; Length 5952;
Best Local Similarity 68.6%; Pred. No. le-257;
Matches 2101; Conservative 0; Mismatches 0; Indels 960; Gaps 1;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTACACCACCTCTGACCCCTACACGGAGCCCAT 60
Db 891 GAGCTATGCCTACACACAGGCTGCTTATGTACACCACCTCTGACCCCTACACGGAGCCCAT 950
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTTGGCAGTTCATTGATGGAGAG 120
Db 951 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTTGGCAGTTCATTGATGGAGAG 1010
QY 121 TGAAGTAAACCTGGACCGTTTATCAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 180
Db 1011 TGAAGTAAACCTGGACCGTTTATCAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 1070
QY 181 TGCTGAGGACACATTGCAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1071 TGCTGAGGACACATTGCAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1130
QY 241 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
Db 1131 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1190
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTATCAGAAGATGAAGA 360
Db 1191 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTATCAGAAGATGAAGA 1250
QY 361 AACTGAAGTACAAGACAGATGAATCTCTTAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
Db 1251 AACTGAAGTACAAGACAGATGAATCTCTTAATTTCAAGATGGGAATGCCTCAGGGTAGC 1310
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 480
Db 1311 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 1370
QY 481 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAGAAAGAAACAAAGGAAATGGAGGAAGA 540
Db 1371 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAGAAAGAAACAAAGGAAATGGAGGAAGA 1430
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCGCAAGTACAACAAATAGGTTGCTTCA 600
Db 1431 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCGCAAGTACAACAAATAGGTTGCTTCA 1490
QY 601 AGAAGATCTAGAACAAGAAACAAGTCAAGGTCAATTTCTCACTCAGATGGTGGTAGT 660
Db 1491 AGAAGATCTAGAACAAGAAACAAGTCAAGGTCAATTTCTCACTCAGATGGTGGTAGT 1550
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAAACAACCTTAAGTATTGGG 720
Db 1551 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAAACAACCTTAAGTATTGGG 1610
QY 721 AGATCGATGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 780
Db 1611 AGATCGATGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 1670
QY 781 CCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTTTAGTGCAATGCTTTTCAGA 840
Db 1671 CCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTTTAGTGCAATGCTTTTCAGA 1730
QY 841 AAAAGAAAGATGCAGTGAACAAGATTCACAACTGGCTTTAAAGATCAAAATGAAATGTT 900

Db 1731 AAAAGAAAGATGCAGTGAACAAGATTACACAACTGGCTTTAAAGATCAAAATGAAATGTT 1790
QY 901 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
Db 1791 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 1850
QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGATAAAGTCAGTGAC 1020
Db 1851 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACACTGAAGATAAAGTCAGTGAC 1910
QY 1021 CCAGAAGACGGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
Db 1911 CCAGAAGACGGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1970
QY 1081 ACTTGAAAAAGAGTACAGCACAG- 1102
Db 1971 ACTTGAAAAAGAGTACAGCACAGGAAACTGAAATAGCAGTTCAAGCTAAACAACCGGATGT 2030
QY 1103 1102
Db 2031 GGAAGAGATTGTTCTAAAGGCGCAGCATTTGTACAAGGAAAAACCAGCCACTCAGCCAGT 2090
QY 1103 1102
Db 2091 GAAGAGGAAGTTAGAAGATCTGAGCTCTGAGTGGGAAGCGGTAAACCGTTTACTTCAAGA 2150
QY 1103 1102
Db 2151 GCTGAGGGCAAGCAGCCCTGACCTAGCTCCTGGACTGACCACCTATTGGAGCCCTCTCTAC 2210
QY 1103 1102
Db 2211 TCAGACTGTTACTCTGTTGACACAACTGTGTTACTTAAGGAAACTGCCCATCTCCAAACT 2270
QY 1103 1102
Db 2271 AGAATGCCATCTTCTTGTATGTTGGAGGTACCTGCTCTGGCAGATTTCAACCGGGCTTG 2330
QY 1103 1102
Db 2331 GACAGAACTACCGACTGGCTTTCTCTGCTTGATCAAGTTATAAATCACAGAGGGTGAT 2390
QY 1103 1102
Db 2391 GGTGGTGACCTTGAGGATATCAACGAGATGATCATCAAGCAGAAGGCAACAATGCAGGA 2450
QY 1103 1102
Db 2451 TTTGGAACAGAGGCGTCCCGAGTTGGAAGAACTCATTAACCGTGCCCCAAAATTTGAAAAA 2510
QY 1103 1102
Db 2511 CAAGACCAGCAATCAAGAGGCTAGAACAAATCATTACGGATCGAATTGAAGAAATTCAGAA 2570
QY 1103 1102
Db 2571 TCAGTGGGATGAAGTACAAGAACACCTTCAGAACCGGAGGCAACAGTTGAATGAAATGTT 2630
QY 1103 1102
Db 2631 AAAGGATTCAACACAAATGGCTGGAAGCTAAGGAAGAAGCTGAGCAGGTCTTAGGACAGGC 2690
QY 1103 1102
Db 2691 CAGAGCCAAGCTTGAGTCATGGAAGGAGGTCCCTATACAGTAGATGCAATCCAAAAGAA 2750
QY 1103 1102
Db 2751 AATCACAGAAACCAACAGATTGGCCAAAGACCTCCGCCAGTGGCAGACAAATGTAGATGT 2810
QY 1103 1102

Db 2811 GGCAAAATGACTTGGCCCTGAAACTTCTCCGGGATTATTCTGCAGATGATACCAGAAAAAGT 2870
QY 1103 1102
Db 2871 CCACATGATAACAGAGAAATATCAATGCCCTCTTGGAGAACATTCATAAAAAGGTGAGTGA 2930
QY 1103 1102
Db 2931 GCGAGAGGCTGCTTTTGGAGAAACTCATAGATTACTGCAACAGTTCCCTCTGGACCTGGA 2990
QY 1141 AAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTAC 1200
Db 2991 AAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTAC 3050
QY 1201 CCGTAAGGAAAGGCTCTAGAAAGACTCCAAAGGAGTAAAAGAGCTGATGAAAACAATGGCA 1260
Db 3051 CCGTAAGGAAAGGCTCTAGAAAGACTCCAAAGGAGTAAAAGAGCTGATGAAAACAATGGCA 3110
QY 1261 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAAACTTGATGAAAACAG 1320
Db 3111 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAAACTTGATGAAAACAG 3170
QY 1321 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTACAAAGACGTTT 1380
Db 3171 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTACAAAGACGTTT 3230
QY 1381 GGATAACATGAACCTCAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGTCCCA 1440
Db 3231 GGATAACATGAACCTCAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGTCCCA 3290
QY 1441 TTTGGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGT 1500
Db 3291 TTTGGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGT 3350
QY 1501 GTGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTTCC 1560
Db 3351 GTGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTTCC 3410
QY 1561 AGCAGTTCAAGAACGACGAAACGATGTACATAGGGCCCTTCAAGAGGGAATTTGAAAACTAAAAGA 1620
Db 3411 AGCAGTTCAAGAACGACGAAACGATGTACATAGGGCCCTTCAAGAGGGAATTTGAAAACTAAAAGA 3470
QY 1621 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGCAGAGCAGCCCTTTGGA 1680
Db 3471 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGCAGAGCAGCCCTTTGGA 3530
QY 1681 AGGACTAGAGAAAACCTCTACAGAGAGCCCGCAGAGAGCTGCCTCTCTGAGGAGAGAGCCAGAA 1740
Db 3531 AGGACTAGAGAAAACCTCTACAGAGAGCCCGCAGAGAGCTGCCTCTCTGAGGAGAGAGCCAGAA 3590
QY 1741 TGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAA 1800
Db 3591 TGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAA 3650
QY 1801 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACT 1860
Db 3651 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACT 3710
QY 1861 TCAAGAGGCCACGGATGAGCTGGAACCTCAAGCTGCGCCAGCTGAGGTGATCAAGGGATC 1920
Db 3711 TCAAGAGGCCACGGATGAGCTGGAACCTCAAGCTGCGCCAGCTGAGGTGATCAAGGGATC 3770
QY 1921 CTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAGATCACCTCGAGAAAAGTCAA 1980
Db 3771 CTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAGATCACCTCGAGAAAAGTCAA 3830
QY 1981 GGCACCTCGAGAGAAAATTGCGCCTCTGAAAGAGAACGCTGAGCCACGTCAATGACCTTGC 2040
Db 3831 GGCACCTCGAGAGAAAATTGCGCCTCTGAAAGAGAACGCTGAGCCACGTCAATGACCTTGC 3890
QY 2041 TCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATTAACCTCAGCACTCTGGAAGA 2100
Db 3891 TCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATTAACCTCAGCACTCTGGAAGA 3950

QY 2101 C 2101

Db 3951 C 3951

RESULT 3

AX114289

LOCUS AX114289 5952 bp DNA linear PAT 11-MAY-2001

DEFINITION Sequence 1 from Patent WO0129243.

ACCESSION AX114289

VERSION AX114289.1 GI:14031259

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Paul,X.L. and Xiao,X.

TITLE Method and vector for producing and transferring trans-spliced peptides

JOURNAL Patent: WO 0129243-A 1 26-APR-2001;

DALHOUSIE UNIVERSITY (CA) ; UNIV. OF PITTSBURGH OF THE COMMONWEALTH SYSTEM OF HIGHER EDUCATION (US)

FEATURES

Location/Qualifiers

1..5952

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

2897..2898

/note="S4 junction site"

3198..3199

/note="S2 junction site"

BASE COUNT 1860 a 1344 c 1410 g 1338 t

ORIGIN

Query Match 53.8%; Score 1131; DB 6; Length 5952;

Best Local Similarity 68.6%; Pred. No. 1e-257;

Matches 2101; Conservative 0; Mismatches 0; Indels 960; Gaps 1;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTACACACCTCTGACCTACACGGAGCCCAATT 60

Db 891 GAGCTATGCCTACACACAGGCTGCTTATGTACACACCTCTGACCTACACGGAGCCCAATT 950

QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120

Db 951 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1010

QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTCGTGGCTTCCTTC 180

Db 1011 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTCGTGGCTTCCTTC 1070

QY 181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240

Db 1071 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1130

QY 241 CCAGTTTCATCTACATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300

Db 1131 CCAGTTTCATCTACATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1190

QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATGGAACAGGAAAATTATCAGAAGATGAAGA 360

Db 1191 TAATATTCTACAATTGGGAAGTAAGCTGATGGAACAGGAAAATTATCAGAAGATGAAGA 1250

QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCATAATTCAAGATGGGAATGCCTCAGGGTAGC 420

Db 1251 AACTGAAGTACAAGAGCAGATGAATCTCCATAATTCAAGATGGGAATGCCTCAGGGTAGC 1310

QY 421 TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480

Db 1311 TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 1370

QY 481 GAAAGAGTTGAATGACTGCTAAACAAAACAGNAGAAAAGAACAGGAAAATGGAGGAAGA 540

Db 1371 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAACAGGAAAATGGAGGAAGA 1430

QY 541 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAAACATAGGTGCTTCA 600

Db 1431 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAAACATAGGTGCTTCA 1490

QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT 660

Db 1491 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT 1550

QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAGAACAACTTAAGGTATTGGG 720

Db 1551 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAGAACAACTTAAGGTATTGGG 1610

QY 721 AGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 780

Db 1611 AGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 1670

QY 781 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTCGCTTTTGTAGTCATGGCTTTTCA 840

Db 1671 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTCGCTTTTGTAGTCATGGCTTTTCA 1730

QY 841 AAAAGAAGATGCAGTGAACAAGATTACACAACAACTGGCTTTAAAGATCAAAATGAAATGTT 900

Db 1731 AAAAGAAGATGCAGTGAACAAGATTACACAACAACTGGCTTTAAAGATCAAAATGAAATGTT 1790

QY 901 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 960

Db 1791 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 1850

QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGAC 1020

Db 1851 GGGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGAC 1910

QY 1021 CCAGAAGACCGAAGCATGGCTGGATAACTTTGCCGGTGTGGGATAATTTAGTCCAAAA 1080

Db 1911 CCAGAAGACCGAAGCATGGCTGGATAACTTTGCCGGTGTGGGATAATTTAGTCCAAAA 1970

QY 1081 ACTTGAAAAGAGTACAGCACAG----- 1102

Db 1971 ACTTGAAAAGAGTACAGCACAGGAAACTGAAATAGCAGTTCAAGCTAAACAAACCGGATGT 2030

QY 1103 ----- 1102

Db 2031 GGAAGAGATTTTGTCTAAAGGGCAGCATTTGTACAAGGAAAACCAGCCACTCAGCCAGT 2090

QY 1103 ----- 1102

Db 2091 GAAAGAGAGTTAGAAAGATCTGAGCTCTGAGTGGAGGGGTAAACCGTTTACTTCAAGA 2150

QY 1103 ----- 1102

Db 2151 GCTGAGGGCAAAGCAGCCCTGACCTAGCTCCTGGACTGACCACCTATTGGAGCCTCTCCTAC 2210

QY 1103 ----- 1102

Db 2211 TCAGACTGTTACTCTGGTGACACAACCTGTGGTTACTAAGGAAACTGCCATCTCCAAACT 2270

QY 1103 ----- 1102

Db 2271 AGAAATGCCATCTTCTTGATGTTGGAGGTACCTGCTCTGGCAGATTCAACCGGGCTTG 2330

QY 1103 ----- 1102

Db 2331 GACAGAACTTACCGACTGGCTTTCTCTGTTGATCAAGTTATATAATCACAGAGGGTGAT 2390

QY 1103 ----- 1102

Db 2391 GGTGGGTGACCTTGAGGATATCAACGAGATGATCATCAAGCAGAGCAACAAATGCAGGA 2450

QY 1103 ----- 1102

Db 2451 TTTGGAACAGAGCGTCCCGAGTTGGAAGAACTCATTACCGCTGCCCAAAATTTGAAAAA 2510
QY 1103 ----- 1102
Db 2511 CAAGACCAGCAATCAAGAGGCTAGAACAATCATTTACGGATCGAATTGAAGAATTCAGAA 2570
QY 1103 ----- 1102
Db 2571 TCAGTGGGATGAAGTACAAGAACACCTTCAGAACCGGAGGCAACAGTTGAATGAATGTT 2630
QY 1103 ----- 1102
Db 2631 AAAGGATTCAACACAATGGCTGGAAGCTAAGGAAGAAGCTGAGCAGGTCTTTAGGACAGGC 2690
QY 1103 ----- 1102
Db 2691 CAGAGCCAAGCTTGAGTCAATGGAAGGAGGGTCCCTATACAGTAGATGCAATCCAAAAGAA 2750
QY 1103 ----- 1102
Db 2751 AATCAGAAACCAAGCAGTTGGCCAAAGACCTCCGCCAGTGGCAGACAAAATGTAGATGT 2810
QY 1103 ----- 1102
Db 2811 GGCAAAATGACTTGGCCCTGAAACTTCTCCGGGATTATTCTCGGAGATATACCAGAAAAGT 2870
QY 1103 ----- 1102
Db 2871 CCACATGATAACAGAGAATATCAATGCCTCTTGGAGAAGCAATTCATAAAGGGTGATGA 2930
QY 1103 ----- 1140
Db 2931 GCGAGAGGCTGCTTTGGAAGAAACTCATAGATTACTGCAACAGTTCCCTCCCTGGACCTGGA 2990
QY 1141 AAAGTTTCTTGCTGGCTTACAGAAGCTGAACAACCTGCCAATGTCTCTACAGGATGTCTAC 1200
Db 2991 AAAGTTTCTTGCTGGCTTACAGAAGCTGAACAACCTGCCAATGTCTCTACAGGATGTCTAC 3050
QY 1201 CCGTAAGGAAAGGCTCCTAGAAGACTCCAGGAGTAAAGAGCTGATGAAACAATGGCA 1260
Db 3051 CCGTAAGGAAAGGCTCCTAGAAGACTCCAGGAGTAAAGAGCTGATGAAACAATGGCA 3110
QY 1261 AGACCTCCAAGGTGAATGAAGCTCACACAGATGTTTATCACAACTGGATGAAACAAG 1320
Db 3111 AGACCTCCAAGGTGAATGAAGCTCACACAGATGTTTATCACAACTGGATGAAACAAG 3170
QY 1321 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGATCAAAAGACGTTT 1380
Db 3171 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGATCAAAAGACGTTT 3230
QY 1381 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGTCCCA 1440
Db 3231 GGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGTCCCA 3290
QY 1441 TTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGT 1500
Db 3291 TTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGT 3350
QY 1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC 1560
Db 3351 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC 3410
QY 1561 AGCAGTTCAGAAAGCAGAACCATGTACATAGGCGCTTCAAGAGGGAATTGAAAACTAAAGA 1620
Db 3411 AGCAGTTCAGAAAGCAGAACCATGTACATAGGCGCTTCAAGAGGGAATTGAAAACTAAAGA 3470
QY 1621 ACCTGTAAATCATGAGTACTCTTGAGACTGTACGAAATATTTCTGACAGAGCAGCTTTGGA 1680
Db 3471 ACCTGTAAATCATGAGTACTCTTGAGACTGTACGAAATATTTCTGACAGAGCAGCTTTGGA 3530
QY 1681 AGGACTAGAGAAACTCTACCAGAGCCAGAGAGCTGCTCCTGAGGAGAGCCCCAGAA 1740
Db 3531 AGGACTAGAGAAACTCTACCAGAGCCAGAGAGCTGCTCCTGAGGAGAGAGCCCCAGAA 3590

QY 1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAAATACTAGTGGGAAAAATTGAA 1800
Db 3591 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAAATACTAGTGGGAAAAATTGAA 3650
QY 1801 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAAGACTCCAGGAACT 1860
Db 3651 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAAGACTCCAGGAACT 3710
QY 1861 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATC 1920
Db 3711 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATC 3770
QY 1921 CTGGCAGCCCGTGGCGGATCTCTCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAAGTCAA 1980
Db 3771 CTGGCAGCCCGTGGCGGATCTCTCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAAGTCAA 3830
QY 1981 GGCACCTTCAGAGAGAAATTGGCGCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGC 2040
Db 3831 GGCACCTTCAGAGAGAAATTGGCGCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGC 3890
QY 2041 TCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATACCTCAGCACCTCTGGAAGA 2100
Db 3891 TCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATACCTCAGCACCTCTGGAAGA 3950
QY 2101 C 2101
Db 3951 C 3951

RESULT 4
LOCUS AX538619 5417 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 39 from Patent WO0229056.
ACCESSION AX538619
VERSION AX538619.1 GI:25271163
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Chamberlain,J.S. and Harper,S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 39 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
source
1. .5417
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"

BASE COUNT 1700 a 1192 c 1182 g 1343 t
ORIGIN

Query Match 52.5%; Score 1103.4; DB 6; Length 5417;
Best Local Similarity 99.9%; Pred. No. 3.6e-251;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGCTATGCCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCCTACACGAGCCCATTT 60
Db 1099 GAGCTATGCCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCCTACACGAGCCCATTT 1158
QY 61 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTTGATGGAGAG 120
Db 1159 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTTGATGGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTT 180
Db 1219 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTT 1278
QY 181 TGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1279 TGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338

| | | | |
|----|------|--|------|
| QY | 241 | CCAGTTTCATCTACTGAGGGGTACATGATGGATTGACAGCCCATCAGGCCGGGTGG | 300 |
| Db | 1339 | CCAGTTTCATCTACTGAGGGGTACATGATGGATTGACAGCCCATCAGGCCGGGTGG | 1398 |
| QY | 301 | TAATATTCTACAATTGGGAAGTAAGCTGATGGAAACAGGAAATATATCAGAAGATGAAGA | 360 |
| Db | 1399 | TAATATTCTACAATTGGGAAGTAAGCTGATGGAAACAGGAAATATATCAGAAGATGAAGA | 1458 |
| QY | 361 | AAGTGAAGTACAAAGCAGATGAATCTCTAAATTCGAATGGGAATGCCTCAGGGTAGC | 420 |
| Db | 1459 | AAGTGAAGTACAAAGCAGATGAATCTCTAAATTCGAATGGGAATGCCTCAGGGTAGC | 1518 |
| QY | 421 | TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT | 480 |
| Db | 1519 | TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT | 1578 |
| QY | 481 | GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAAGAAAAATGGAGGAAGA | 540 |
| Db | 1579 | GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAAGAAAAATGGAGGAAGA | 1638 |
| QY | 541 | GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACATAAAGGTGCTTCA | 600 |
| Db | 1639 | GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACATAAAGGTGCTTCA | 1698 |
| QY | 601 | AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCAATGCTGGGTAGT | 660 |
| Db | 1699 | AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCAATGCTGGGTAGT | 1758 |
| QY | 661 | TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG | 720 |
| Db | 1759 | TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG | 1818 |
| QY | 721 | AGATCGATGGCAACATCTGTAGATGGACAGAGACCGCTGGGTCTTTTACAAGACAT | 780 |
| Db | 1819 | AGATCGATGGCAACATCTGTAGATGGACAGAGACCGCTGGGTCTTTTACAAGACAT | 1878 |
| QY | 781 | CCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTATGTCATGGCTTTTCAGA | 840 |
| Db | 1879 | CCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTATGTCATGGCTTTTCAGA | 1938 |
| QY | 841 | AAAAGAAGATGCAGTGAACAAAGATTTCACACAACAGTGGCTTTAAAGATCAAAATGAAATGTT | 900 |
| Db | 1939 | AAAAGAAGATGCAGTGAACAAAGATTTCACACAACAGTGGCTTTAAAGATCAAAATGAAATGTT | 1998 |
| QY | 901 | ATCAAGTCTTCAAAACTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT | 960 |
| Db | 1999 | ATCAAGTCTTCAAAACTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT | 2058 |
| QY | 961 | GGGCAAACTGTATTCTCAAAACAAGATCTTCTTCAACACCTGAAGAAATAAGTCAGTGAC | 1020 |
| Db | 2059 | GGGCAAACTGTATTCTCAAAACAAGATCTTCTTCAACACCTGAAGAAATAAGTCAGTGAC | 2118 |
| QY | 1021 | CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA | 1080 |
| Db | 2119 | CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA | 2178 |
| QY | 1081 | ACTTGAAAAGAGTACAGCAGACT 1105 | |
| Db | 2179 | ACTTGAAAAGAGTACAGCAGACT 2203 | |

| | | | | | | | |
|------------|------------------------------------|-------------|----------|----------|-----|--------|-----------------|
| RESULT 5 | AX538624 | LOCUS | AX538624 | 11443 bp | DNA | linear | PAT 23-NOV-2002 |
| DEFINITION | Sequence 44 from Patent WO0229056. | | | | | | |
| ACCESSION | AX538624 | | | | | | |
| VERSION | AX538624.1 | GI:25271175 | | | | | |
| KEYWORDS | synthetic construct | | | | | | |
| SOURCE | synthetic construct | | | | | | |
| ORGANISM | artificial sequences. | | | | | | |
| REFERENCE | 1 | | | | | | |

AUTHORS Chamberlain,J.S. and Harper,S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 44 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES Location/Qualifiers
source 1. .11443
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"
BASE COUNT 3707 a 2339 c 2502 g 2895 t
ORIGIN

| | | | | | |
|-----------------------|------|--|---------------------|-----------|---------------|
| Query Match | | 52.5%; | Score 1103.4; | DB 6; | Length 11443; |
| Best Local Similarity | | 99.9%; | Pred. No. 3.6e-251; | | |
| Matches 1104; | | Conservative 0; | Mismatches 1; | Indels 0; | Gaps 0; |
| QY | 1 | GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAT | 60 | | |
| Db | 1099 | GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAT | 1158 | | |
| QY | 61 | TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG | 120 | | |
| Db | 1159 | TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG | 1218 | | |
| QY | 121 | TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC | 180 | | |
| Db | 1219 | TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC | 1278 | | |
| QY | 181 | TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA | 240 | | |
| Db | 1279 | TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA | 1338 | | |
| QY | 241 | CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG | 300 | | |
| Db | 1339 | CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG | 1398 | | |
| QY | 301 | TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTTATCAGAAGATGAAGA | 360 | | |
| Db | 1399 | TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTTATCAGAAGATGAAGA | 1458 | | |
| QY | 361 | AACTGAAGTACAAGACAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC | 420 | | |
| Db | 1459 | AACTGAAGTACAAGACAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC | 1518 | | |
| QY | 421 | TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT | 480 | | |
| Db | 1519 | TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT | 1578 | | |
| QY | 481 | GAAAGAGTTGAATGACTGGCTTAAACAAAAACAGAAAGAAACAAGAAAAATGGAGGAAGA | 540 | | |
| Db | 1579 | GAAAGAGTTGAATGACTGGCTTAAACAAAAACAGAAAGAAACAAGAAAAATGGAGGAAGA | 1638 | | |
| QY | 541 | GCCTCTTGGACCTGATCTTGAAGACCTTAAAAACGCCAAGTACAAACAATAAGGTGCTTCA | 600 | | |
| Db | 1639 | GCCTCTTGGACCTGATCTTGAAGACCTTAAAAACGCCAAGTACAAACAATAAGGTGCTTCA | 1698 | | |
| QY | 601 | AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCAATGCTGGTGGTAGT | 660 | | |
| Db | 1699 | AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCAATGCTGGTGGTAGT | 1758 | | |
| QY | 661 | TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG | 720 | | |
| Db | 1759 | TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG | 1818 | | |
| QY | 721 | AGATCGATGGGCAACATCTGTAGATGGACAGAAACCGCTGGGTCTTTTACAAGACAT | 780 | | |
| Db | 1819 | AGATCGATGGGCAACATCTGTAGATGGACAGAAACCGCTGGGTCTTTTACAAGACAT | 1878 | | |
| QY | 781 | CCTTCTCAAATGGCAACGCTTACTGAAGAAACAGTGCCTTTTTTAGTCATGGCTTTTCAGA | 840 | | |
| Db | 1879 | CCTTCTCAAATGGCAACGCTTACTGAAGAAACAGTGCCTTTTTTAGTCATGGCTTTTCAGA | 1938 | | |

QY 841 AAAAGAGATGAGTGAACAAAGATTACACAACTGGGTTTAAAGATCAAAATGAAATGTT 900
Db 1939 AAAAGAGATGAGTGAACAAAGATTACACAACTGGGTTTAAAGATCAAAATGAAATGTT 1998
QY 901 ATCAAGTCTTCAAAAACCTGGCGGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
Db 1999 ATCAAGTCTTCAAAAACCTGGCGGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2058
QY 961 GGGCAAACTGTATTCACTCAAAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db 2059 GGGCAAACTGTATTCACTCAAAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2118
QY 1021 CCAGAGACGGGAAGCATGGCTGGATAAATTGGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
Db 2119 CCAGAGACGGGAAGCATGGCTGGATAAATTGGCCCGGTGTTGGGATAATTTAGTCCAAAA 2178
QY 1081 ACTTGAAGAGTACAGCACAGACT 1105
Db 2179 ACTTGAAGAGTACAGCACAGATT 2203
RESULT 6
AX538627 12057 bp DNA linear PAT 23-NOV-2002
LOCUS AX538627
DEFINITION Sequence 47 from Patent WO0229056.
ACCESSION AX538627
VERSION AX538627.1 GI:25271181
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE
1 Chamberlain,J.S. and Harper,S.Q.
AUTHORS Mini-dystrophin nucleic acid and peptide sequences
TITLE Patent: WO 0229056-A 47 11-APR-2002;
JOURNAL THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
1. 12057
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"
BASE COUNT 4020 a 2456 c 2790 g 2791 t
ORIGIN
Query Match 52.5%; Score 1103.4; DB 6; Length 12057;
Best Local Similarity 99.9%; Pred.No.3.6e-251;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCCTACACGGAGCCCAT 60
Db 1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCCTACACGGAGCCCAT 1158
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAACACAAGTCATTTGGCAGTTCATTGATGGAGAG 120
Db 1159 TCCTTCACAGCATTTGGAAGCTCCTGAACACAAGTCATTTGGCAGTTCATTGATGGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAAGTATATCGTGGCTTCCTTC 180
Db 1219 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAAGTATATCGTGGCTTCCTTC 1278
QY 181 TGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAATGATGTGAAGTGGTGAAGA 240
Db 1279 TGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAATGATGTGAAGTGGTGAAGA 1338
QY 241 CCAGTTTCATCTCAGGGGTACATGATGATGATGATGATGATGATGATGATGATGATGATG 300
Db 1339 CCAGTTTCATCTCAGGGGTACATGATGATGATGATGATGATGATGATGATGATGATG 1398
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
Db 1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1458

QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 1518
QY 421 TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 480
Db 1519 TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 1578
QY 481 GAAAGAGTTGAATGACTGGCTTAAACAAACAAAGAAAGAAACAAAGAAATGGAGGAAGA 540
Db 1579 GAAAGAGTTGAATGACTGGCTTAAACAAACAAAGAAAGAAACAAAGAAATGGAGGAAGA 1638
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACACATAAGTGTCTTCA 600
Db 1639 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACACATAAGTGTCTTCA 1698
QY 601 AGAAGATCTAGAACAAAGCAAGTCAAGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 660
Db 1699 AGAAGATCTAGAACAAAGCAAGTCAAGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 1758
QY 661 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAAACAACTTAAGGTATTGGG 720
Db 1759 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAAACAACTTAAGGTATTGGG 1818
QY 721 AGATCGATGGGCAAAACATCTCTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 780
Db 1819 AGATCGATGGGCAAAACATCTCTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 1878
QY 781 CCTTCTCAAAATGGCAACGCTTCTTACTGAAGAACAGTGCCTTTTGTGATGGCTTTTCAGA 840
Db 1879 CCTTCTCAAAATGGCAACGCTTCTTACTGAAGAACAGTGCCTTTTGTGATGGCTTTTCAGA 1938
QY 841 AAAAGAGATGCAGTGAACAAAGATTACACAACTGGCTTTAAAGATCAAAATGAAATGTT 900
Db 1939 AAAAGAGATGCAGTGAACAAAGATTACACAACTGGCTTTAAAGATCAAAATGAAATGTT 1998
QY 901 ATCAAGTCTTCAAAAACCTGGCGGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
Db 1999 ATCAAGTCTTCAAAAACCTGGCGGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2058
QY 961 GGGCAAACTGTATTCACTCAAAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db 2059 GGGCAAACTGTATTCACTCAAAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2118
QY 1021 CCAGAGACGGGAAGCATGGCTGGATAAATTGGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
Db 2119 CCAGAGACGGGAAGCATGGCTGGATAAATTGGCCCGGTGTTGGGATAATTTAGTCCAAAA 2178
QY 1081 ACTTGAAGAGTACAGCACAGACT 1105
Db 2179 ACTTGAAGAGTACAGCACAGATT 2203
RESULT 7
AX409637 13957 bp DNA linear PAT 14-JUN-2002
LOCUS AX409637
DEFINITION Sequence 2284 from Patent WO0229103.
ACCESSION AX409637
VERSION AX409637.1 GI:21442342
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
AUTHORS Gene expression profiles in liver cancer
TITLE Patent: WO 0229103-A 2284 11-APR-2002;
JOURNAL GENE LOGIC INC (US)
FEATURES
Location/Qualifiers
1. 13957
/organism="Homo sapiens"
/mol_type="genomic DNA"

/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. M18533"
BASE COUNT 4602 a 2781 c 3122 g 3452 t
ORIGIN

Query Match 52.5%; Score 1103.4; DB 6; Length 13957;
Best Local Similarity 99.9%; Pred. No. 3.6e-251;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGGAGCCCAT 60
Db 1099 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGGAGCCCAT 1158

QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 120
Db 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAAGTATATCGTGGCTTCTTC 180
Db 1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAAGTATATCGTGGCTTCTTC 1278
QY 181 TGCTGAGGACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1279 TGCTGAGGACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338

QY 241 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
Db 1339 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1398
QY 301 TAATATTCTACAAATTGGGAAGTAAGTCTCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 360
Db 1399 TAATATTCTACAAATTGGGAAGTAAGTCTCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1458

QY 361 AACTGAAGTACAGAGCAGATGAATCTCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAGAGCAGATGAATCTCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1518
QY 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 480
Db 1519 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 1578

QY 481 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAACAAAGGAAATGGAGGAAGA 540
Db 1579 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAACAAAGGAAATGGAGGAAGA 1638

QY 541 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAAACAATAAGGTGCTTCA 600
Db 1639 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAAACAATAAGGTGCTTCA 1698

QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGTAGT 660
Db 1699 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGTAGT 1758

QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 720
Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 1818

QY 721 AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGCTGGGTCTTTTACAAGACAT 780
Db 1819 AGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGCTGGGTCTTTTACAAGACAT 1878

QY 781 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTTTAGTCATGGCTTTTCAGA 840
Db 1879 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTTTAGTCATGGCTTTTCAGA 1938

QY 841 AAAAGAAGATGCAGTGAACAAGATTCCACAACTGGCTTTTAAAGATCAAAATGAAATGTT 900
Db 1939 AAAAGAAGATGCAGTGAACAAGATTCCACAACTGGCTTTTAAAGATCAAAATGAAATGTT 1998

QY 901 ATCAAGTCTTCAAAAATGGCCCGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
Db 1999 ATCAAGTCTTCAAAAATGGCCCGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2058

QY 961 GGGCAAACTGTATTCACTCAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db 2059 GGGCAAACTGTATTCACTCAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2118
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTTAGTCCAAAA 1080
Db 2119 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTTAGTCCAAAA 2178
QY 1081 ACTTGAAAAGAGTACAGCACAGACT 1105
Db 2179 ACTTGAAAAGAGTACAGCACAGATT 2203

RESULT 8
AX538581
LOCUS AX538581 13957 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 1 from Patent WO0229056.
ACCESSION AX538581
VERSION AX538581.1 GI:25271086
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 1 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

FEATURES
source Location/Qualifiers
1..13957
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 4602 a 2781 c 3122 g 3452 t
ORIGIN

Query Match 52.5%; Score 1103.4; DB 6; Length 13957;
Best Local Similarity 99.9%; Pred. No. 3.6e-251;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGGAGCCCAT 60
Db 1099 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGGAGCCCAT 1158

QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 120
Db 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 1218

QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAAGTATATCGTGGCTTCTTTC 180
Db 1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAAGTATATCGTGGCTTCTTTC 1278

QY 181 TGCTGAGGACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1279 TGCTGAGGACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338

QY 241 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
Db 1339 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1398

QY 301 TAATATTCTACAAATTGGGAAGTAAGTCAATTTGGAACAGGAAATTTATCAGAAGATGAAGA 360
Db 1399 TAATATTCTACAAATTGGGAAGTAAGTCAATTTGGAACAGGAAATTTATCAGAAGATGAAGA 1458

QY 361 AACTGAAGTACAGAGCAGATGAATCTCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAGAGCAGATGAATCTCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1518

QY 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 480
Db 1519 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 1578

| | | | |
|------------|--|--|--------|
| QY | 481 | GAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAAGAACAAAGAAATGGAGGAAGA | 540 |
| Db | 1579 | GAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAAGAACAAAGAAATGGAGGAAGA | 1638 |
| QY | 541 | GCCTCTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAAACAATAAGGTGCTTCA | 600 |
| Db | 1639 | GCCTCTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAAACAATAAGGTGCTTCA | 1698 |
| QY | 601 | AGAGATCTAGAACAAAGAACAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT | 660 |
| Db | 1699 | AGAGATCTAGAACAAAGAACAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT | 1758 |
| QY | 661 | TGATGAATCTAGTGGAGATACCGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG | 720 |
| Db | 1759 | TGATGAATCTAGTGGAGATACCGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG | 1818 |
| QY | 721 | AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGTCTCTTTTACAAGACAT | 780 |
| Db | 1819 | AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGTCTCTTTTACAAGACAT | 1878 |
| QY | 781 | CCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTGTGCAATGGCTTTTCAGA | 840 |
| Db | 1879 | CCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTGTGCAATGGCTTTTCAGA | 1938 |
| QY | 841 | AAAAGAAGATGCAGTGAACAAGATTCAACAACCTGGCTTTAAAGATCAAATGAAATGTT | 900 |
| Db | 1939 | AAAAGAAGATGCAGTGAACAAGATTCAACAACCTGGCTTTAAAGATCAAATGAAATGTT | 1998 |
| QY | 901 | ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT | 960 |
| Db | 1999 | ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT | 2058 |
| QY | 961 | GGGCAAACTGATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC | 1020 |
| Db | 2059 | GGGCAAACTGATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC | 2118 |
| QY | 1021 | CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA | 1080 |
| Db | 2119 | CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA | 2178 |
| QY | 1081 | ACTTGAAAAGAGTACAGCACAGACT | 1105 |
| Db | 2179 | ACTTGAAAAGAGTACAGCACAGATT | 2203 |
| RESULT 9 | | | |
| HUMDYS | | 13957 bp | linear |
| LOCUS | | | |
| DEFINITION | Homo sapiens dystrophin (DMD) mRNA, complete cds. | | |
| ACCESSION | M18533 M17154 M18026 M20250 | | |
| VERSION | M18533.1 GI:181856 | | |
| KEYWORDS | | | |
| SOURCE | Homo sapiens (human) | | |
| ORGANISM | Homo sapiens | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | |
| REFERENCE | 1 (bases 1 to 1699) | | |
| AUTHORS | Koenig,M., Hoffman,E.P., Bertelson,C.J., Monaco,A.P., Feener,C. and Kunkel,L.M. | | |
| TITLE | Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals | | |
| JOURNAL | Cell 50 (3), 509-517 (1987) | | |
| MEDLINE | 87273512 | | |
| PUBMED | 3607877 | | |
| REFERENCE | 2 (bases 1678 to 3830) | | |
| AUTHORS | Hoffman,E.P., Monaco,A.P., Feener,C.C. and Kunkel,L.M. | | |
| TITLE | Conservation of the Duchenne muscular dystrophy gene in mice and humans | | |
| JOURNAL | Science 238 (4825), 347-350 (1987) | | |
| MEDLINE | 88018015 | | |
| PUBMED | 3659917 | | |
| REFERENCE | 3 (bases 1 to 13957) | | |

| | |
|----------|---|
| AUTHORS | Koenig,M., Monaco,A.P. and Kunkel,L.M. |
| TITLE | The complete sequence of dystrophin predicts a rod-shaped cytoskeletal protein |
| JOURNAL | Cell 53 (2), 219-226 (1988) |
| MEDLINE | 88194521 |
| PUBMED | 3282674 |
| COMMENT | On May 25, 2000 this sequence version replaced gi:340693. Draft entry and computer-readable sequence kindly provided by M.Koenig, 01-APR-1988 The severity of muscular dystrophy is determined by the size of the deleted DNA segment. Deletions found in different patients were from positions 302-2200, 473-1168, 1691-1810, and 1169-3011. |
| FEATURES | Location/Qualifiers |
| source | 1..13957 |
| | /organism="Homo sapiens" |
| | /mol_type="mRNA" |
| | /db_xref="taxon:9606" |
| | /chromosome="X" |
| | /map="Xp21.3-p21.1" |
| | /tissue_type="muscle" |
| | /dev_stage="fetus" |
| | /note="G00-119-850" |
| gene | 1..13957 |
| | /gene="DMD" |
| CDS | 209..11266 |
| | /gene="DMD" |
| | /codon_start=1 |
| | /product="dystrophin" |
| | /protein_id="AAA53189.1" |
| | /db_xref="GI:181857" |
| | /translation="MLWEEVEDCYEREDVQKKTFTKWVNAQFSKFGKOHENLFSDL QDGRLLDLLEGLTGKLPKEKGSTRVHALNNVNNKALRVLQNNVDLVNIGSTDIVDG NHKLTGLIWNILHWQVKNVMKNIMAGLQQTNSKILLSWVRQSTRYPQVNVINF TSWSDGLALNALIHSRPPDLFDWNSVVCQOSATQRLAHAFNIARYQLGIEKLLDPEDV DTTYPDKKSILMYITSLFQVLPQOVSIEAIQEVEMLPKPKVTKEEHFOLHHQMHYSQ QITVSLAQGYERTSSPKRFSYATQAAVYTTSDPTRSPFPSSHLEAPDKSFGSSL MESEVNLDRYQTALIEEVLSSLSAEDTLQAQGEISNDVEVVKDQFHTHEGYMMDLTAH QGRVGNILQGLKLTGTGLSEDETEVEQMNLLNSRWECLRVASMEKSNLHRVLM DLQNKLELNDLWLTKEERTKMEEEPGLDLEDLKRQVQKHVQLQEDLEQEBQVRVN SLTHMVVVDESSEGHATAALEEQLVKLGDRWANI CRWTEDRWVLLQDILLKQRLTE EQCLFSAWLSEKEDAVNKIHTTGFKDQNEMLSSLOKLAVLKADLEKKKQSMGKLYSLK QDLSTLTKNSVTOKTEAWLDFNFCWDLNVQKLEKSTAQISQAVTTTQPSLTQTVM ETVTVTTREQILVXHAQEELPPPPPKKQKITVDSEIRKRLDVIDITELHSWTRSEA VLQSPFAIFRKEGNFSDLKEKVNAIEREKAEKFRKLQDASRQAQALVEQMVNEGVNA DSIIKQASEQLNSRWTFEQQLSERLNWLEYQNNIIAFYNQLQLEQMTTAAENWLKIQ PTPPSEPTAIKSQKICKDEVNRLSGLQPOIERLKIQSIALKEKGQGPMLDADFVAF TNHFKQVSDVQAREKELQTIIDTLPMPRYQETMSAIRTWVQOSETKLSIPQLSVTDY EIMBQRLGELQALQSSLQEQSGLYLSTTVKEMSKAPSEISRKYQSEFEIEGRWK KLSSQLVEHCQKLEEQMNKLRKIQNHIIQTLKKMAEVDVFLKEEWPALGDSIELKKQL KQCRLLVSDIQTIOPSLNSVNEGQKIKNEAEPEFASRLTELKELTNOWDMCQQVY ARKEALKGGLEKTVSLQKLSMEHMTQAESEYLERDFFEYKTPDELQKXAVEMKRAK EEAQQKEAKVKLLTESVNSVIAQAPPVAQEAALKKELETLTNYQWLCTRLNGKCKTLE EVWACWHELLSYLEKANKWNEVEFKLTENIPGAAEISEVLDSELENMRHSDNP NQIRILAQTLTDGGVMDLINEELETFNSRWRELHEEAVRRQKLEQSIQSAQETESK LHLIQESLTFIDKQLAAYIADKVDAQMPQEAQKIQSDLTSHSIEISLEEMKHNQKKEA AQRVLSQIDVAQKQLQVSMKFRLLFQKPA NFELQLQESKMILDEVKMHLPALETKSVE QEVVQSQNLHCNVNLYKSLSEVKSEVEMVIKTGRQIVQKKQNTENPKELDERVTALKLHY NELGAKVTERKQQLKCLKLSRKMRKEMNVLTEWLAATDMELTKRSAVEGMPNSLDSSE VAWGKATQKEIEKQKVHLKSI TEVGEALKTVLGKETLVEDKLSLLNSNWIATVTSRAE EWLNLLEYQKHMETFDQNVDHITKWI IQADTLLDSEKKKPPQKEDVLKRLKAEI LND IRPKVDS TRDQAANL MANRGDHCRLVEPQISELNHFAAISHR IKTGKASIPUKLE QFNSDIQKLLLEPLEAEIQGVNKEEDFNKMDNENEGTVKELLQRGDNLQORITDER KREEIKIQKLLQKHNALDRLSQRKKALEISHQYKQRQADLLKCLDDIEKKL ASLPEPRDERKIKEIDRELQKKKEELNAVRRQAEGISEDGAAMAVEPTQIQLSKWRE IESKFAQFRRLNFAQIHTVREETMMVTMEDPLEISYVPSYTLTEITHVSQALLVEEQ LLNAPDLCAKDFEDLFKQESLKNIKDSLQSSSGRIDIHSKKTAALQASATPVERVKL QEALSQIDFQWEKVNKYKDRQDRFSVEKWRFRHYDIKIFNQLWTEAEQFLRKTQI PENWEHAKYKWLKELQDGIGQRTVVRTLNATGEEIIQSSKTDASILQEKLGSLNL RWQEVCKQLSDRKKRLEEQKNI LSEFQDLNFEVLWEADNIASIPLEPGKEQQLKE KLEQVKLLVEELPLRQGIKQLNETGGPVLVSAPISEEDQKLENKLTQTNLQWIKVS RALPEKQGEIEAQIKDLGQLEKKLEDLEEQLNLHLWLSPIRNQLEIYNQNOEGPFD VQETETAVQAQKQPDVEEILSKQHLKKEPATQVVKRLEDLSSEKAVNRLLQELRA KQPD LAPGLTTIGASPTQTVTILVTQFVVTKETAISKLEMPSSIMLEVPALADFNRAWT |

ELTDWLSLLDQVKSQRVMVGDDLEDINEMIIKQKATMODLEQRRPQLEELITAAQNLK
NKTSNQEARIIITDRIERI QNWDEVQEHQNRQOLNEMLKDSQWLEAKEEAEQVL
GOARAKLESWKEGPTVDAIQKII TETKOLAKDLRQWQTNVDVANDIALKLLRDYSAD
DTRKVMITENINASWRSIHKRVSEREAALBETHRLLLQFPPLDLEKFLAWLTEAETTA
NVLDQATRKERLLEDKQWELMKOWDLOGIEAHTDVYHNLNDSQKILRSLEGS
DAVLLQRRLDNNFNKSELKKSINIRSHLEASSDQWKRLLHLSQELLVWLQKDDDEL
SRQAPIGGDFPAVQKQNDVHRAFKRELKTFEPMVIMSTLETVRIIFLTPQPLEGLEKLYQ
EPRELPPERAQNVTRLLRKQAEVENTEWELNHLHSAADWQKIDETLERLHQBELQEAOT
ELDLKLRQAEVYIKGSQWPGVGLLIDSLQDHLEKVKALRGEIAPLKENVSHVNDLQARL
TTLGIQLSPYNLSTLEDNLRWKLQVAVEDRVQOLHEHRDFGPASQHFLLSTSVQGP
WERAI SPNKVPYVINHETQTTCWDHPKMTELYQSLADLNNVRFSAYRTAMKLLRLOKA
LCDDLLLSAACDALDQNLKQNDQFMDILOI INCLTIIYDRLEQEHNNLVNPLCVD
MCLNWLNNVYDTGRTGIRVLSFKTGIISLCKAHLKEDKYRYLFKQVASSTFGCDQRRLL
GLLLHDSIQIPRQLGEVASFSGSNIIEPSVSCFQFANNKRYFAKHPRMGYLPVQTVLEGNMET
VWLPVLRVAAAEATAKHAQKCNICKECPPIIGFRYRSLKHFNVDICQSCFFSGRVAKGH
KWHYPMVEYCTPTTSGEDVRDFAKVLKNKFRFKRYFAKHPRMGYLPVQTVLEGNMET
PVTLINFVPVDSAPASSPOLSHDDTHSRIEHYASRLAEMENSNGSYLNDISIPNESID
DEHLLIOHYCQSLNQDPSLQSPRPAQILISLESEERLELERILADLEEENRNLOAEY
DRLKQOHEHKGLSPLPSPPEMPTSPQSPRDAELIAEAKLLRQHKRLEARMQILEDH
NKQLESQHLRLRQLLEPOAEAKVNGTTVSSPSTLQSRDSSQPMMLRVVGSQTSDSM
GEEDLLSPFPQDTSTGLEEVNEQLNNSFPSSRGRNTPGKPMREDTM"

BASE COUNT 4602 a 2781 c 3122 g 3452 t
ORIGIN

Query Match 52.5%; Score 1103.4; DB 9; Length 13957;
Best Local Similarity 99.9%; Pred. No. 3.6e-251;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2Y 1 GAGCTATGCCTACACACAGGCTGCTTATGTACCAACCTCTGACCCCTACACGGAGCCCAT 60
2b 1099 GAGCTATGCCTACACACAGGCTGCTTATGTACCAACCTCTGACCCCTACACGGAGCCCAT 1158
2Y 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
2b 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1218
2Y 121 TGAAGTAAACCTGGACCGGTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
2b 1219 TGAAGTAAACCTGGACCGGTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1278
2Y 181 TGCTGAGGACACATTTGAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
2b 1279 TGCTGAGGACACATTTGAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
2Y 241 CCAGTTTCTACTCATGAGGGGTACATGATGGATTTGCACAGCCCATCAGGCGCGGTTGG 300
2b 1339 CCAGTTTCTACTCATGAGGGGTACATGATGGATTTGCACAGCCCATCAGGCGCGGTTGG 1398
2Y 301 TAATATTCTACAATTTGGGAAGTAAAGCTGATTTGGAACAGGAGAAATTTATCAGAAGATGAAGA 360
2b 1399 TAATATTCTACAATTTGGGAAGTAAAGCTGATTTGGAACAGGAGAAATTTATCAGAAGATGAAGA 1458
2Y 361 AACTGAAGTACAAGACGAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
2b 1459 AACTGAAGTACAAGACGAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1518
2Y 421 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480
2b 1519 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 1578
2Y 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAGGAAATGGAGGAAGA 540
2b 1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAGGAAATGGAGGAAGA 1638
2Y 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACAGCCCAAGTACAAACAATAAGGTGCTTCA 600
2b 1639 GCCTCTTGGACCTGATCTTGAAGACCTTAAACAGCCCAAGTACAAACAATAAGGTGCTTCA 1698
2Y 601 AGAAGATCTAGAAACAAGAAACAAGTCAGGGTCAATTTCTCTACTCACATGGTGGTGTAGT 660
2b 1699 AGAAGATCTAGAAACAAGAAACAAGTCAGGGTCAATTTCTCTACTCACATGGTGGTGTAGT 1758
2Y 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAAACAACCTTAAGGTATTGGG 720

Db 1759 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1818
QY 721 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 780
Db 1819 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 1878
QY 781 CCTTCTCAAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTATAGTGCATGGCTTTTACA 840
Db 1879 CCTTCTCAAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTATAGTGCATGGCTTTTACA 1938
QY 841 AAAAGAGATGCAGTGAACCAAGATTACACAACTGGCTTTTAAAGATCAAAAATGAAATGTT 900
Db 1939 AAAAGAGATGCAGTGAACCAAGATTACACAACTGGCTTTTAAAGATCAAAAATGAAATGTT 1998
QY 901 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
Db 1999 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2058
QY 961 GGGCAAACTGTATTCACTCAAAACAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db 2059 GGGCAAACTGTATTCACTCAAAACAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2118
QY 1021 CCAGAACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
Db 2119 CCAGAACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 2178
QY 1081 ACTTGAAAAGAGTACAGCACAGACT 1105
Db 2179 ACTTGAAAAGAGTACAGCACAGATT 2203

RESULT 10
AR220819
LOCUS AR220819 13977 bp DNA linear PAT 26-SEP-2002
DEFINITION Sequence 60 from patent US 6426186.
ACCESSION AR220819
VERSION AR220819.1 GI:23327696
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 13977)
AUTHORS Jones,K.A., Volkmut,W. and Walker,M.G.
TITLE Bone remodeling Genes
JOURNAL Patent: US 6426186-A 60 30-JUL-2002;
FEATURES Location/Qualifiers
source
1. .13977
/organism="unknown"
BASE COUNT 4596 a 2765 c 3120 g 3453 t 43 others
ORIGIN

Query Match 52.5%; Score 1103.4; DB 6; Length 13977;
Best Local Similarity 99.9%; Pred. No. 3.6e-251;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTACCAACCTCTGACCCCTACACGGAGCCCAT 60
Db 1099 GAGCTATGCCTACACACAGGCTGCTTATGTACCAACCTCTGACCCCTACACGGAGCCCAT 1158
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
Db 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGGTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
Db 1219 TGAAGTAAACCTGGACCGGTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1278
QY 181 TGCTGAGGACACATTTGAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1279 TGCTGAGGACACATTTGAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
QY 241 CCAGTTTCTACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCGCGGTTGG 300

| | | | |
|----|------|---|------|
| Db | 1339 | CCAGTTTCATACTCATGAGGGGTACATGATGGATTGTGACAGCCCATCAGGCCGGTGG | 1398 |
| QY | 301 | TAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA | 360 |
| Db | 1399 | TAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA | 1458 |
| QY | 361 | AACTGAAGTACAAGACGAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC | 420 |
| Db | 1459 | AACTGAAGTACAAGACGAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC | 1518 |
| QY | 421 | TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAA | 480 |
| Db | 1519 | TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAA | 1578 |
| QY | 481 | GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAGGAAAATGGAGGAAGA | 540 |
| Db | 1579 | GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAGGAAAATGGAGGAAGA | 1638 |
| QY | 541 | GCCTCTTGGACCTGATCTTGAGAGACCTAAACGCCAAGTACAACAACATAAGTGCTTCA | 600 |
| Db | 1639 | GCCTCTTGGACCTGATCTTGAGAGACCTAAACGCCAAGTACAACAACATAAGTGCTTCA | 1698 |
| QY | 601 | AGAAGATCTAGAACAAAGTCAAGTCAAGGTCAATTCTCTCACTCACATGGTGGTAGT | 660 |
| Db | 1699 | AGAAGATCTAGAACAAAGTCAAGTCAAGGTCAATTCTCTCACTCACATGGTGGTAGT | 1758 |
| QY | 661 | TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAAACAACTTAAGTATGGG | 720 |
| Db | 1759 | TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAAACAACTTAAGTATGGG | 1818 |
| QY | 721 | AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT | 780 |
| Db | 1819 | AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT | 1878 |
| QY | 781 | CCTTCTCAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTAGTGCAATGGCTTTCAGA | 840 |
| Db | 1879 | CCTTCTCAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTAGTGCAATGGCTTTCAGA | 1938 |
| QY | 841 | AAAAGAAGATGCAGTGAAACAAGATTACACAACTGGCTTTTAAAGATCAAAATGAATGTT | 900 |
| Db | 1939 | AAAAGAAGATGCAGTGAAACAAGATTACACAACTGGCTTTTAAAGATCAAAATGAATGTT | 1998 |
| QY | 901 | ATCAAGTCTTCAAAAACTGGCCGCTTTTAAAGCGGATCTAGAAAAAGAAAAGCAATCCAT | 960 |
| Db | 1999 | ATCAAGTCTTCAAAAACTGGCCGCTTTTAAAGCGGATCTAGAAAAAGAAAAGCAATCCAT | 2058 |
| QY | 961 | GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC | 1020 |
| Db | 2059 | GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC | 2118 |
| QY | 1021 | CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTTAGTCCAAAA | 1080 |
| Db | 2119 | CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTTAGTCCAAAA | 2178 |
| QY | 1081 | ACTTGAAAAAGATACAGCACAGACT 1105 | |
| Db | 2179 | ACTTGAAAAAGATACAGCACAGATT 2203 | |

RESULT 11
HSDMDR Human mRNA for dystrophin. 12446 bp mRNA linear PRI 12-SEP-1993
LOCUS X14298
DEFINITION Dmd gene; Duchenne muscular dystrophy; dystrophin.
ACCESSION X14298
VERSION X14298.1 GI:30845
KEYWORDS Dmd gene; Duchenne muscular dystrophy; dystrophin.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 12446)
AUTHORS Rosenthal,A., Speer,A., Billwitz,H., Cross,G.S., Forrest,S.M. and

Davies,K.E.
Two human cDNA molecules coding for the Duchenne muscular dystrophy (DMD) locus are highly homologous
Nucleic Acids Res. 17 (13), 5391 (1989)
89345106
2668885
2 (bases 1 to 12446)
Rosenthal,A.
Direct Submission
Submitted (09-FEB-1989) Rosenthal A., Akademie der Wissenschaften der DDR, Zentralinstitut fuer Molekularbiologie, Robert-Roessle Str.10, 1115 Berlin Buch, DDR
see also M18533 and M20250 for Dmd seqs.; discrepancies compared to M18533 cDNA were located at x14298 pos. 496, 1772, 1965, 2449, 3687, 4229, 4504, 5075, 5332, 5630 and 7194.
Location/Qualifiers
1. 12446
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X Chromosomal, Xp21."
/tissue_type="muscle"
/dev stage="fetal and adult."
99. 11156
/note="dystrophin (AA 1 - 3685)"
/codon_start=1
/protein_id="CAA32479.1"
/db_xref="GI:30846"
/db_xref="SWISS-PROT:P11532"
/translation="MLWEEVEDCYEREDVQKKTFTKWNNAQFSKFGKQHIENLFSDL QDGRRLDLLLEGLTGQKLPKEGSTRVHALNNVNKALRVLQNNVNDLVNIGSTDIVDG NHKLTGLIWNIIILHWQVKNVMKNIMAGLOPTNSEKILLSWVRQSTRNYPQVNVINFT TSWDGLALNALIHSHRPDLFDWNSVVCOSATORLEHAFNARIYQLGIEKLLDPEDV DTTYPKKSILMYTISLFOVLPQQVSIEAIEVEMLPRPPKVTKEEHFQLHHQMHYSQ QITVSLAQQYERTSSPKPRFKSYAYTOAAYVTSDTPSPFSPQSHLEAPEDKSPGSSL MESEVNLDRYQYTALEEVLWSLLSAEDTLQAQGEISNDVEVKDQFTHGEGYMDLTAH QGRVGNIIQLGSKLIGTKLSEDEETEVOQMNLNLSRWECLRVASMEKQSNLHRLVM DLQNKLKLNDMLTKTEERTRKMEEPGLGPDLEDLKRQVQOHKVLQEDLEQEOVRVN SLTHMVVWDESSGDHATAALEEQKVLGDRWANI CRWTEDRWVLLDILLKQWRLTE EQCLFSAWLSEKEDAVNKIHTTGFKDQDNEMLSLQKLAVLKADLEKKQSMGKLYSIK QDLSTLKNKSVTKTEAWLDNFARCDNLVQKLEKSTAQISOAVTTQPSLTQTIVM ETVTVTTRREQILVKHAQEELPPPPQKKRQITVDSEIRKRLDVIDTELHSWITRSEA VLQSPFAIFRKEGNFSDLKEKVNAIEREKAEKFRKLQDASRSGOALVEQMVNEGVNA DSIOQASEQLNSRWTFECQLLSERLNWLEYQNNIIAFYNQLQQLQEQMTTAENWLKIQ KLSQLVEHCQKLEEQMNKLKIQNHIIQTLKKMAEVDVFLKEWPAFGDSEILKKQL KQCRLLVSDIQTIQPSLNSVNEGQKIKNBAEPEPASRLTEKELNLTQWDHMCQQVY ARKEALKGLEKTVSLQKDLSEHMEWMTQAEYEYLFEDFEYKTPDELQKAFEEEMKRAK EEAQOKEAKVKLLTESVNSVIAQAPPAQEAALKKELETLTTNYQWLCTRLNGKCKTLE EFWACWHELLSYLEKANKWLNVEVEFKLKTENIPGGAAEISEVLDSLENLMRHSEDNF NQIRILAQTLTDGVDDELINEELETNFSNRWELHEEAVRRQKLEQSIQSAQETENS LHIIQESLTFIDKQLAAAYIADKVDAQAQPOEAQIKQSDLTSHIEISLEEMKKHNOGKEA AQRVLSQIDVAQKQLQDVSMKFRLLFQKPAQFQRLQESKMILDEVKMLPALETKSVE QEVVQSQLNHCVNLYKSLSEVKSEVEMVIKTGRQIVQKQKQENPKELDERVTALKHY NELGAKVTERKQQLKLEKCLSRKMRKEMNVLTEWLAATDMELTKRSAGEGMPNLDSE VAWGKATQKEIEKQVHLKSI TEVGEALKTVLGKETLVEDKLSLLNSNWIATVTSRAE EWLNLILLEQKHMETFDQNVDDHITKWI IQADTLDESEKKKPOQKEDVLKRLKAEI LDPKVDSTRDQAANLMAHGDHCKRLVEPQISELNHRFAAISHRIKTKGASIPKLELE QFNSDIQKLEPLEAEIQQGVNLKBEEDFNKMNEDNEGTVKVELLQQRNLOQRITDER KSEEEKIKQOLLQTKNALKDLRSQRRKKALEISHQWYQYKQADDLKCLDDIEKKL ASLPEPRDERKIKEIDRELQKKEELNARVROAEGLEDGAMAVEPTQIQLSKRWRE IESKPAQFRRLNFAQIHTVREETMMVMTEDMPLEISYVPSYLTETHTVSOALLEVEQ ILNAPDLCAQDFEDLFQOESLKNIKDSLQSSGRFIDIHSHKTAALQASATPVERVKL QEALSQLDFQWEKVNKMYKDRQGRFDRSVKWRFRHYDIKFNQWLTAEOFLRKTQI PENWEHAKYKWLKELQDGIGQRTVVRTLNATGEEIIQSSKTDASILEKLGSLNL RWQEVCKQLSDRKKRLEEKNILSEFORDLNEFVLWEADNIASIPLEPGKQOOLKE KLEQVKLLVEELPLRQGLKQLNETGGPVLVSAPISPEEQDKLENKLTQTNLOWIKVS RALPEKQGEIEAQIKDLGQLEKKLEDLEEQLNHLNLSPIRNQLEIYNQPNQGPFD VKETETIAVQAKQPDVVEILSKGHLYKEKPATQPKVRLKLEDLSSEWKA VNRLLQELRA KQPD LAPGLTTIGASPTQVTLVTPQVVTKETATSKLEMPSSLMLEVPALADFNRAWT ELTDWLSLLDQVIKSRQVMVGDLEDINEMIIKQKATMQDLEQRRRPOLEELITAAQNLK

NKTSNOEARTIITDRIERIONQWDEVQEHQNRQQLNEMLKDSTOWLEAKEEAEQVL
GQARAKLESWEKGYTVDAIOKKITETKQAKDLROWQTNVDVANDLALKLLRDYSAD
DTRKVMITENINASWRSIHKRVSEREALEETHRLLOQFPDLEKFLAWLTEAETTA
NVLDATRKERLLEDSKGVKELMKQWDOQEIEAHTDVYHNLDSNQKILRSLEGS
DAVLLQRRLLDNMFKWSELKKSLSNIRSHLEASDDQWKRHLHLSQELLVWLQKDEL
SROAPIGGDFPAVQKQNDVHRAFKRELKTKBPVIMSTLETVTFTEQPLEGLEKLYQ
EPRELPPERAQNVTRLRQKAEVNTWEKLNHLSADWQKIDETLERLOELQEA
ELDLKROAEVIGKSWQPVGDLLIDSLQDHLKLEKALRGEIAPLKNVSHVNDLARQ
TTLGQIQLSPYNLSTLEDNTRWKLLOVAVEDVRQOLHEHRDFGASQHFLLSTSVQGP
WERALSPNKVPYVINHETQTTCDWHPKMTLYQSLADLNVRFSAYRTAMKLRLOKA
LCLDLLSLSAACDALDQHNLKQNDQPMIDILQINCLTIIYDRLEQEHNNLVNPLCVD
MCLNWLNLVYDTGRTRIRVLSFKTGIISLCKAHLKDYRVLFXOVASSTGFCDDORRL
GLLLHDSIQIPRQGEVASFGGSNIEPSVRSQFQFANKKPEIEALFLDWMRLPQSM
VWLPVLRVAAAEATAKHQAKNICKECPILIGFRYRSLKHENYDICOQCFSSGRVAKGH
KMHYPMVEYCTPTTSGEDVRDFAKVLKNKFKTKRYFAKHPRMGVLPVQTVLEGNMET
PVTLINFVPVDSAPASSPQLSHDDTHSRIEHYASRLAEMENSNGSVLNDISI
DEHLLIQHYCQSLNODSPLSQPRSPAQILISLESEERGELEIRLADLEENRNLOAEY
DRLKQOHEHKGLSPLSPPEMPTSPQSPRDAELIAEAKLLRQHKRLEARMQILEDH
NKQLESQHLRLQLLEPOAEAKVNGTIVSSPSTSLQRSDSSQPNMLLRVVGSQTSDSM
GEEDLLSPQDSTSTGLEVMEQLNNSFPSSRGENTPGKPMREDTM"
BASE COUNT 4135 a 2524 c 2876 g 2911 t
ORIGIN

Query Match 52.4%; Score 1100.2; DB 9; Length 12446;
Best Local Similarity 99.7%; Pred. No. 2.1e-250;
Matches 1102; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

2y 1 GAGCTATGCCTACACACAGGCTGCTTATGTCCACCCTCTGACCCCTACACGGAGCCCAT 60
|||
3b 989 GAGCTATGCCTACACACAGGCTGCTTATGTCCACCCTCTGACCCCTACACGGAGCCCAT 1048
|||
2y 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
|||
3b 1049 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1108
|||
2y 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 180
|||
Db 1109 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 1168
|||
Qy 181 TGCTGAGGACACATTTGCAAGCACACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
|||
Db 1169 TGCTGAGGACACATTTGCAAGCACACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1228
|||
Qy 241 CCAGTTTCTACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGCCGGGTGG 300
|||
Db 1229 CCAGTTTCTACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGCCGGGTGG 1288
|||
Qy 301 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
|||
Db 1289 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1348
|||
Qy 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
|||
Db 1349 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1408
|||
Qy 421 TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 480
|||
Db 1409 TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 1468
|||
Qy 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAACAGGAAAATGGAGGAAGA 540
|||
Db 1469 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAACAGGAAAATGGAGGAAGA 1528
|||
Qy 541 GCCTCTTGGACCTGATCTTTGAAGACCTAAAAACGCCAAGTACAAACAATAGGTGCTTCA 600
|||
Db 1529 GCCTCTTGGACCTGATCTTTGAAGACCTAAAAACGCCAAGTACAAACAATAGGTGCTTCA 1588
|||
Qy 601 AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTTCTCTACTCACATGGTGGTGTAGT 660
|||
Db 1589 AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTTCTCTACTCACATGGTGGTGTAGT 1648
|||
Qy 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
|||

Db 1649 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1708
Qy 721 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGTCTTTTACAGACAT 780
|||
Db 1709 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGTCTTTTACAGACAT 1768
|||
Qy 781 CCTTCTCAAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTAGTGCAATGGCTTTTCA 840
|||
Db 1769 CCTGCTCAAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTAGTGCAATGGCTTTTCA 1828
|||
Qy 841 AAAAGAAGATGCAGTGAACAAGATTACACAACCTGGCTTTAAAGATCAAAATGAAATGTT 900
|||
Db 1829 AAAAGAAGATGCAGTGAACAAGATTACACAACCTGGCTTTAAAGATCAAAATGAAATGTT 1888
|||
Qy 901 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
|||
Db 1889 ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 1948
|||
Qy 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTTCAACACTGAAGATAAGTCAGTGAC 1020
|||
Db 1949 GGGCAAACTGTATTCAATCAAAACAAGATCTTCTTTTCAACACTGAAGATAAGTCAGTGAC 2008
|||
Qy 1021 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
|||
Db 2009 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 2068
|||
Qy 1081 ACTTGAAAAGAGTACAGCACAGACT 1105
Db 2069 ACTTGAAAAGAGTACAGCACAGATT 2093
|||
RESULT 12
AX538620
LOCUS AX538620 5339 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 40 from Patent WO0229056.
ACCESSION AX538620
VERSION AX538620.1 GI:25271166
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 40 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
source
1. .5339
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"
BASE COUNT 1638 a 1191 c 1187 g 1323 t
ORIGIN
Query Match 47.5%; Score 997; DB 6; Length 5339;
Best Local Similarity 100.0%; Pred. No. 6.9e-226;
Matches 997; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1105 TCATAGATTACTGCAACAGTTCCCTGGACCTGGAAAAGTTTCTTGCTGGCTTACAGA 1164
|||
Db 1546 TCATAGATTACTGCAACAGTTCCCTGGACCTGGAAAAGTTTCTTGCTGGCTTACAGA 1605
|||
Qy 1165 AGCTGAAACAACCTGCCAATGTCTTACAGGATGTACCCCGTAAGGAAAGGCTCCTTAGAAGA 1224
|||
Db 1606 AGCTGAAACAACCTGCCAATGTCTTACAGGATGTACCCCGTAAGGAAAGGCTCCTTAGAAGA 1665
|||
Qy 1225 CTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGC 1284
|||
Db 1666 CTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGC 1725
|||
Qy 1285 TCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAATAATCCTGAGATCCCTGGA 1344
|||

Db 1726 TCACACAGATGTTTATCACAACTGGATGAAAAACAGCCAAAAATCCTGAGATCCCTGGA 1785
QY 1345 AGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAG 1404
Db 1786 AGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAG 1845
QY 1405 TGAACCTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTCTTGACCAGTG 1464
Db 1846 TGAACCTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTCTTGACCAGTG 1905
QY 1465 GAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGA 1524
Db 1906 GAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGA 1965
QY 1525 ATTAAGCCGGCAGGCACCTTATTTGGAGGCGACTTTCCAGCAGTTTCAGAAAGCAGACGATGT 1584
Db 1966 ATTAAGCCGGCAGGCACCTTATTTGGAGGCGACTTTCCAGCAGTTTCAGAAAGCAGACGATGT 2025
QY 1585 ACATAGGGCCTTCAAGAGGGGAATTGAAACCTAAAGAACCTGTAAATCATGAGTACTTTGA 1644
Db 2026 ACATAGGGCCTTCAAGAGGGGAATTGAAACCTAAAGAACCTGTAAATCATGAGTACTTTGA 2085
QY 1645 GACTGTACGAATATTTCTGCACAGCAGCCCTTTGGAAGGACTAGAGAACTCTACCAGGA 1704
Db 2086 GACTGTACGAATATTTCTGCACAGCAGCCCTTTGGAAGGACTAGAGAACTCTACCAGGA 2145
QY 1705 GCCACAGAGCTGCCTCCTCGAGAGAGAGCCCCAGAAATGTCACTCGGCTTCTACGAAAGCA 1764
Db 2146 GCCACAGAGCTGCCTCCTCGAGAGAGAGCCCCAGAAATGTCACTCGGCTTCTACGAAAGCA 2205
QY 1765 GGCTGAGGAGTCAATACTAGTGGGAAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAG 1824
Db 2206 GGCTGAGGAGTCAATACTAGTGGGAAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAG 2265
QY 1825 AAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGA 1884
Db 2266 AAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGA 2325
QY 1885 CCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCT 1944
Db 2326 CCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCT 2385
QY 1945 CATTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCATTCGAGGAGAAATTGCGCC 2004
Db 2386 CATTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCATTCGAGGAGAAATTGCGCC 2445
QY 2005 TCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGTTACCACCTTTGGGCAT 2064
Db 2446 TCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGTTACCACCTTTGGGCAT 2505
QY 2065 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGAC 2101
Db 2506 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGAC 2542

RESULT 13
AX538621
LOCUS AX538621 5462 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 41 from Patent WO0229056.
ACCESSION AX538621
VERSION AX538621.1 GI:25271168
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 41 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
location/Qualifiers
FEATURES
source 1..5462
/organism="synthetic construct"

/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"
BASE COUNT 1668 a 1225 c 1212 g 1357 t
ORIGIN
Query Match 47.4%; Score 996; DB 6; Length 5462;
Best Local Similarity 100.0%; Pred. No. 1.2e-225;
Matches 996; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1106 CATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAAGTTTCTTGCTCGCTTACAGAA 1165
Db 1670 CATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAAGTTTCTTGCTCGCTTACAGAA 1729
QY 1166 GGTGAAAAAAGTGTCCATGTCTCTACAGGATGCTACCCGTAAAGAAAGGCTCCTAGAAGAC 1225
Db 1730 GGTGAAAAAAGTGTCCATGTCTCTACAGGATGCTACCCGTAAAGAAAGGCTCCTAGAAGAC 1789
QY 1226 TCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAATTTGAAGCT 1285
Db 1790 TCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAATTTGAAGCT 1849
QY 1286 CACACAGATGTTTATCAACAACCTGGATGAAACAAGCCAAAAAATCCTGAGATCCCTGGAA 1345
Db 1850 CACACAGATGTTTATCAACAACCTGGATGAAACAAGCCAAAAAATCCTGAGATCCCTGGAA 1909
QY 1346 GGTTCGGATGATGCAGTCTCTGTTTCAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGT 1405
Db 1910 GGTTCGGATGATGCAGTCTCTGTTTCAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGT 1969
QY 1406 GAACCTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTTCTGACCAGTGG 1465
Db 1970 GAACCTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTTCTGACCAGTGG 2029
QY 1466 AAGCGTCTGCACCTTTCTCTGAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAA 1525
Db 2030 AAGCGTCTGCACCTTTCTCTGAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAA 2089
QY 1526 TTAAGCCGGCAGGCACCTATTGAGGCGGACTTTCAGCAGTTCAGAAAGCAGAACGATGTA 1585
Db 2090 TTAAGCCGGCAGGCACCTATTGAGGCGGACTTTCAGCAGTTCAGAAAGCAGAACGATGTA 2149
QY 1586 CATAGGGCCTTCAAGAGGGAAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAG 1645
Db 2150 CATAGGGCCTTCAAGAGGGAAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAG 2209
QY 1646 ACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGAG 1705
Db 2210 ACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGAG 2269
QY 1706 CCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCAG 1765
Db 2270 CCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCAG 2329
QY 1766 GCTGAGGAGTCAATACTGAGTGGGAAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGA 1825
Db 2330 GCTGAGGAGTCAATACTGAGTGGGAAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGA 2389
QY 1826 AAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGAC 1885
Db 2390 AAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGAC 2449
QY 1886 CTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCTC 1945
Db 2450 CTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCTC 2509
QY 1946 ATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCT 2005
Db 2510 ATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCT 2569
QY 2006 CTGAAAGAGAACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATT 2065
Db 2570 CTGAAAGAGAACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATT 2629

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model
Run on: February 1, 2004, 10:14:55 ; Search time 4634.58 Seconds
(without alignments)
11483.979 Million cell updates/sec

Title: US-09-845-416-10_COPY_1800_3100
Perfect score: 1301
Sequence: 1 cgactttccagcagtttcaga.....ccggagctgcttccaatttg 1301

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB | ID | Description |
|------------|--------|---------------|--------|----|-----------|---------------------|
| 1 | 1301 | 100.0 | 5339 | 6 | AX538620 | AX538620 Sequence |
| 2 | 1301 | 100.0 | 5462 | 6 | AX538621 | AX538621 Sequence |
| 3 | 1301 | 100.0 | 5952 | 6 | AR304538 | AR304538 Sequence |
| 4 | 1301 | 100.0 | 5952 | 6 | AX114289 | AX114289 Sequence |
| 5 | 1301 | 100.0 | 8689 | 6 | AX538622 | AX538622 Sequence |
| 6 | 1301 | 100.0 | 11443 | 6 | AX538624 | AX538624 Sequence |
| 7 | 1301 | 100.0 | 12057 | 6 | AX538627 | AX538627 Sequence |
| 8 | 1301 | 100.0 | 12446 | 9 | HSDMDR | X14298 Human mRNA |
| 9 | 1301 | 100.0 | 13957 | 6 | AX409637 | AX409637 Sequence |
| 10 | 1301 | 100.0 | 13957 | 6 | AX538581 | AX538581 Sequence |
| 11 | 1301 | 100.0 | 13957 | 9 | HUMDYS | M18533 Homo sapien |
| 12 | 1290 | 99.2 | 13977 | 6 | AR220819 | AR220819 Sequence |
| 13 | 1154.4 | 88.7 | 13887 | 4 | AF070485 | AF070485 Canis fam |
| 14 | 1135.2 | 87.3 | 13815 | 6 | AX306153 | AX306153 Sequence |
| 15 | 1135.2 | 87.3 | 13815 | 6 | AX538582 | AX538582 Sequence |
| 16 | 1135.2 | 87.3 | 13815 | 10 | MUSDYSA | M68859 Mouse dystro |
| 17 | 1135.2 | 87.3 | 19307 | 6 | AR093392 | AR093392 Sequence |
| 18 | 1135.2 | 87.3 | 19307 | 6 | AR142592 | AR142592 Sequence |
| 19 | 1085.2 | 83.4 | 4402 | 6 | E30219 | E30219 Shortened d |
| 20 | 1016.6 | 78.1 | 5417 | 6 | AX538619 | AX538619 Sequence |
| 21 | 904.8 | 69.5 | 13575 | 5 | GGDYS | X13369 Chicken mRN |
| 22 | 886 | 68.1 | 4402 | 6 | E30220 | E30220 Shortened d |
| 23 | 750.4 | 57.7 | 4075 | 6 | E30221 | E30221 Shortened d |
| 24 | 746.2 | 57.4 | 3747 | 6 | E30218 | E30218 Shortened d |
| 25 | 722.4 | 55.5 | 2654 | 5 | FSCDYSTRO | M37645 Torpedo cal |
| 26 | 692.6 | 53.2 | 3163 | 6 | E30223 | E30223 Shortened d |
| 27 | 608.2 | 46.7 | 3521 | 5 | AF339031 | AF339031 Danio rer |
| 28 | 584 | 44.9 | 4658 | 9 | HUMDMDX | M92650 Human Duche |
| 29 | 584 | 44.9 | 4658 | 9 | BC028720 | BC028720 Homo sapi |
| 30 | 570.4 | 43.8 | 10705 | 10 | RNAJ2967 | AJ002967 Rattus no |
| 31 | 567.2 | 43.6 | 3161 | 10 | MMGUTRPH | X83506 M.musculus |
| 32 | 567.2 | 43.6 | 11096 | 6 | AX538584 | AX538584 Sequence |
| 33 | 567.2 | 43.6 | 11096 | 10 | MMY12229 | Y12229 M.musculus |
| 34 | 553.8 | 42.6 | 6045 | 6 | A63605 | A63605 Sequence 7 |
| 35 | 553.8 | 42.6 | 6045 | 6 | AR281528 | AR281528 Sequence |
| 36 | 553.8 | 42.6 | 6059 | 6 | AX107972 | AX107972 Sequence |
| 37 | 553.8 | 42.6 | 10302 | 6 | AX538583 | AX538583 Sequence |
| 38 | 553.8 | 42.6 | 10302 | 9 | HSMUPS | X69086 H.sapiens m |
| 39 | 553.8 | 42.6 | 10320 | 6 | A63607 | A63607 Sequence 9 |
| 40 | 553.8 | 42.6 | 10320 | 6 | AR281529 | AR281529 Sequence |
| 41 | 551 | 42.4 | 3499 | 9 | HSU43519 | U43519 Human dyst |
| 42 | 551 | 42.4 | 5106 | 6 | AX552248 | AX552248 Sequence |
| 43 | 546.8 | 42.0 | 3073 | 10 | AF195788 | AF195788 Rattus no |
| 44 | 546.8 | 42.0 | 3140 | 10 | AF195787 | AF195787 Rattus no |
| 45 | 529.8 | 40.7 | 1240 | 10 | RNAPDY3 | X69767 R.norvegicu |

ALIGNMENTS

| | | | | |
|------------|--|-------------|-----|--------|
| RESULT 1 | | | | |
| AX538620 | AX538620 | 5339 bp | DNA | linear |
| LOCUS | Sequence 40 from Patent WO0229056. | | | |
| DEFINITION | AX538620 | | | |
| ACCESSION | AX538620.1 | GI:25271166 | | |
| VERSION | | | | |
| KEYWORDS | synthetic construct | | | |
| SOURCE | synthetic construct | | | |
| ORGANISM | artificial sequences. | | | |
| REFERENCE | 1 | | | |
| AUTHORS | Chamberlain,J.S. and Harper,S.Q. | | | |
| TITLE | Mini-dystrophin nucleic acid and peptide sequences | | | |
| JOURNAL | Patent: WO 0229056-A 40 11-APR-2002; | | | |
| | THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US) | | | |

FEATURES
source
Location/Qualifiers
1. 5339
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"
BASE COUNT 1638 a 1191 c 1187 g 1323 t
ORIGIN
Query Match 100.0%; Score 1301; DB 6; Length 5339;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
Db 1993 CGACTTTCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 2052
QY 61 AACTAAAGAACTGTAAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db 2053 AACTAAAGAACTGTAAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCA 2112
QY 121 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG 180
Db 2113 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG 2172
QY 181 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 240
Db 2173 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 2232
QY 241 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTCAAAAGACT 300
Db 2233 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTCAAAAGACT 2292
QY 301 CCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 360
Db 2293 CCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 2352
QY 361 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 420
Db 2353 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 2412
QY 421 GAAAGTCAAGGCACTTCGAGGAGAAATTCGGCCTCTGAAAGAGAAACGTGAGCCACGTCAA 480
Db 2413 GAAAGTCAAGGCACTTCGAGGAGAAATTCGGCCTCTGAAAGAGAAACGTGAGCCACGTCAA 2472
QY 481 TGACCTTGCTGCCAGCTTACCACCTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCAC 540
Db 2473 TGACCTTGCTGCCAGCTTACCACCTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCAC 2532
QY 541 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGCTCGAGGACCGAGTCAG 600
Db 2533 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGCTCGAGGACCGAGTCAG 2592
QY 601 GCAGCTGATGAAGCCCAAGGCACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
Db 2593 GCAGCTGATGAAGCCCAAGGCACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 2652
QY 661 TGTCCAGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACCA 720
Db 2653 TGTCCAGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACCA 2712
QY 721 CGAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCGAGTCTTTAGC 780
Db 2713 CGAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCGAGTCTTTAGC 2772
QY 781 TGACCTGAATAATGTCAAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
Db 2773 TGACCTGAATAATGTCAAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 2832
QY 841 GAAGGCCCTTTGCTTGATCTCTTGAGCCTGTGCAGCTGCATGTGATGCCTTGGACCAGCA 900
Db 2833 GAAGGCCCTTTGCTTGATCTCTTGAGCCTGTGCAGCTGCATGTGATGCCTTGGACCAGCA 2892

QY 901 CAACCTCAAGCAAAATGACAGCCCATGGATATCTCTGCAGATTATTAAATTGTTGACCAC 960
Db 2893 CAACCTCAAGCAAAATGACAGCCCATGGATATCTCTGCAGATTATTAAATTGTTGACCAC 2952
QY 961 TATTATGACCGCCTGGAGCAAGAGCACAACTTTGGTCAACGTCCCTCTCTGCGTGA 1020
Db 2953 TATTATGACCGCCTGGAGCAAGAGCACAACTTTGGTCAACGTCCCTCTCTGCGTGA 3012
QY 1021 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1080
Db 3013 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 3072
QY 1081 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
Db 3073 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 3132
QY 1141 ATACCTTTTCAAGCAAGTGCAGAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 1200
Db 3133 ATACCTTTTCAAGCAAGTGCAGAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 3192
QY 1201 CCTTCTGCATGATTTCTATCCAAATTTCCAAGACAGTTGGTGAAGTTGCATCCTTTGGGG 1260
Db 3193 CCTTCTGCATGATTTCTATCCAAATTTCCAAGACAGTTGGTGAAGTTGCATCCTTTGGGG 3252
QY 1261 CAGTAAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAATTG 1301
Db 3253 CAGTAAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAATTG 3293
RESULT 2
AX538621
LOCUS AX538621 5462 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 41 from Patent WO0229056.
ACCESSION AX538621
VERSION AX538621.1 GI:25271168
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 41 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
Source
1. 5462
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"
BASE COUNT 1668 a 1225 c 1212 g 1357 t
ORIGIN
Query Match 100.0%; Score 1301; DB 6; Length 5462;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
Db 2116 CGACTTTCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 2175
QY 61 AACTAAAGAACTGTAAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db 2176 AACTAAAGAACTGTAAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCA 2235
QY 121 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG 180
Db 2236 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG 2295
QY 181 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 240
Db 2296 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 2355

QY 241 AAAATTGAACCTGCACTCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGAAGACT 300
Db 2356 AAAATTGAACCTGCACTCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGAAGACT 2415
QY 301 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGAT 360
Db 2416 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGAT 2475
QY 361 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCAAGATCACCTCGA 420
Db 2476 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCAAGATCACCTCGA 2535
QY 421 GAAAGTCAAGGCACCTCGAGGAGAAATTGGCCTCTGAAAGAGAACTGAGCCACGTCAA 480
Db 2536 GAAAGTCAAGGCACCTCGAGGAGAAATTGGCCTCTGAAAGAGAACTGAGCCACGTCAA 2595
QY 481 TGACCTTGCTCGCCAGCTTACCACCTTTGGSCATTGAGCTCTCACCGTATAACCTCAGCAC 540
Db 2596 TGACCTTGCTCGCCAGCTTACCACCTTTGGSCATTGAGCTCTCACCGTATAACCTCAGCAC 2655
QY 541 TCTGGAAGACCTGAACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 600
Db 2656 TCTGGAAGACCTGAACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 2715
QY 601 GCAGCTGCATGAAGCCACAGGGAACCTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTC 660
Db 2716 GCAGCTGCATGAAGCCACAGGGAACCTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTC 2775
QY 661 TGTCCAGGGTCCCTGGAGAGAGCCCATCTGCCAAACAAAGTGCCCTACTATATCAACCA 720
Db 2776 TGTCCAGGGTCCCTGGAGAGAGCCCATCTGCCAAACAAAGTGCCCTACTATATCAACCA 2835
QY 721 CGAGACTCAAAACAATTGCTGGGACCATCCCAAAATGACAGAGCTTACCAGTCTTTAGC 780
Db 2836 CGAGACTCAAAACAATTGCTGGGACCATCCCAAAATGACAGAGCTTACCAGTCTTTAGC 2895
QY 781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCA 840
Db 2896 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCA 2955
QY 841 GAAAGCCCTTTGCTTGGATCTTTGAGCCTGTCAGCTGCATGTGATGCTTGGACCAAGCA 900
Db 2956 GAAAGCCCTTTGCTTGGATCTTTGAGCCTGTCAGCTGCATGTGATGCTTGGACCAAGCA 3015
QY 901 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCCAC 960
Db 3016 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCCAC 3075
QY 961 TATTTATGACCGCTCGAGCAAGAGCACAACTTTGTTATGATACGGGACGAACAGGAGGATCCGTGT 1020
Db 3076 TATTTATGACCGCTCGAGCAAGAGCACAACTTTGTTATGATACGGGACGAACAGGAGGATCCGTGT 3135
QY 1021 TATGTGTCTGAACCTGGTCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 1080
Db 3136 TATGTGTCTGAACCTGGTCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 3195
QY 1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGAAGACAAAGTACAG 1140
Db 3196 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGAAGACAAAGTACAG 3255
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGAGCGCAGGCTGGGCCT 1200
Db 3256 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGAGCGCAGGCTGGGCCT 3315
QY 1201 CCTTCTGCATGATTCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGTCATCTTTGGGG 1260
Db 3316 CCTTCTGCATGATTCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGTCATCTTTGGGG 3375
QY 1261 CAGTAAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAATTG 1301
Db 3376 CAGTAAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAATTG 3416

RESULT 3
AR304538
LOCUS AR304538 5952 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 1 from patent US 6544786.
ACCESSION AR304538
VERSION AR304538.1 GI:31693691
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 5952)
AUTHORS Xiao,X. and Liu,P.X.
.TITLE Method and vector for producing and transferring trans-spliced peptides
JOURNAL Patent: US 6544786-A 1 08-APR-2003;
FEATURES Location/Qualifiers
source 1..5952
BASE COUNT 1860 a 1344 c 1410 g 1338 t
ORIGIN
Query Match 100.0%; Score 1301; DB 6; Length 5952;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
Db 3402 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 3461
QY 61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db 3462 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 3521
QY 121 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCTGAGGAGAG 180
Db 3522 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCTGAGGAGAG 3581
QY 181 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 240
Db 3582 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 3641
QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGAAGACT 300
Db 3642 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGAAGACT 3701
QY 301 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 360
Db 3702 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 3761
QY 361 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 420
Db 3762 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 3821
QY 421 GAAAGTCAAGGCACCTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTCGAGCCACGTCAA 480
Db 3822 GAAAGTCAAGGCACCTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTCGAGCCACGTCAA 3881
QY 481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTGAGCTCTCACCGTATAACCTCAGCAC 540
Db 3882 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTGAGCTCTCACCGTATAACCTCAGCAC 3941
QY 541 TCTGGAAGACCTGAACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 600
Db 3942 TCTGGAAGACCTGAACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 4001
QY 601 GCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTCCACGTC 660
Db 4002 GCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTCCACGTC 4061
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCAAAACAAAGTGCCCTACTATATCAACCA 720
Db 4062 TGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCAAAACAAAGTGCCCTACTATATCAACCA 4121

| | | | |
|-----------------------|------|---|------|
| QY | 721 | CGAGACTCAAACTTGCTGGGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGC | 780 |
| Db | 4122 | CGAGACTCAAACTTGCTGGGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGC | 4181 |
| QY | 781 | TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA | 840 |
| Db | 4182 | TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA | 4241 |
| QY | 841 | GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCCTTGGACCAGCA | 900 |
| Db | 4242 | GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCCTTGGACCAGCA | 4301 |
| QY | 901 | CAACCTCAAGCAAAATGACCGCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC | 960 |
| Db | 4302 | CAACCTCAAGCAAAATGACCGCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC | 4361 |
| QY | 961 | TATTTATGACCGCTGGAGCAAGAGCAACAATTGTTCAACGTCCCTCTCTGCGTGGGA | 1020 |
| Db | 4362 | TATTTATGACCGCTGGAGCAAGAGCAACAATTGTTCAACGTCCCTCTCTGCGTGGGA | 4421 |
| QY | 1021 | TATGTGTCTGAATGCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGT | 1080 |
| Db | 4422 | TATGTGTCTGAATGCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGT | 4481 |
| QY | 1081 | CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGACAAGTACAG | 1140 |
| Db | 4482 | CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGACAAGTACAG | 4541 |
| QY | 1141 | ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGCCT | 1200 |
| Db | 4542 | ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGCCT | 4601 |
| QY | 1201 | CCTTCTGCATGATTCTATCCAAATTCCTGAGAGTGGGTGAAGTTGCATCCTTTGGGGG | 1260 |
| Db | 4602 | CCTTCTGCATGATTCTATCCAAATTCCTGAGAGTGGGTGAAGTTGCATCCTTTGGGGG | 4661 |
| QY | 1261 | CAGTAACATTGAGCCAACTGTCGGAGCTGCTTCCAATTG | 1301 |
| Db | 4662 | CAGTAACATTGAGCCAACTGTCGGAGCTGCTTCCAATTG | 4702 |
| RESULT 4 | | | |
| AX114289 | | | |
| LOCUS | | | |
| DEFINITION | | | |
| ACCESSION | | | |
| VERSION | | | |
| KEYWORDS | | | |
| SOURCE | | | |
| ORGANISM | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| FEATURES | | | |
| source | | | |
| misc_feature | | | |
| misc_feature | | | |
| BASE COUNT | | | |
| ORIGIN | | | |
| Query Match | | | |
| Best Local Similarity | | | |

| | | | |
|---|------|--|------|
| Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | |
| QY | 1 | CGACTTTCCAGCAGTTTCAGAAAGCAGAAACGATGTACATAGGGCCTTCAAGAGGGAAATTGAA | 60 |
| Db | 3402 | CGACTTTCCAGCAGTTTCAGAAAGCAGAAACGATGTACATAGGGCCTTCAAGAGGGAAATTGAA | 3461 |
| QY | 61 | AACTAAAGAACCTGTAAATCATAGTACTCTTTGAGACTGTACGAATATTTCTTGACAGAGCA | 120 |
| Db | 3462 | AACTAAAGAACCTGTAAATCATAGTACTCTTTGAGACTGTACGAATATTTCTTGACAGAGCA | 3521 |
| QY | 121 | GCCTTTTGAAGGACTAGAGAACTCTACCAAGAGCCAGAGAGCTGCCTCCTGAGGAGAG | 180 |
| Db | 3522 | GCCTTTTGAAGGACTAGAGAACTCTACCAAGAGCCAGAGAGCTGCCTCCTGAGGAGAG | 3581 |
| QY | 181 | AGCCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCAATACTGAGTGGGA | 240 |
| Db | 3582 | AGCCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCAATACTGAGTGGGA | 3641 |
| QY | 241 | AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT | 300 |
| Db | 3642 | AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT | 3701 |
| QY | 301 | CCAGGAACTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT | 360 |
| Db | 3702 | CCAGGAACTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT | 3761 |
| QY | 361 | CAAGGGATCCTGGCAGCCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA | 420 |
| Db | 3762 | CAAGGGATCCTGGCAGCCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA | 3821 |
| QY | 421 | GAAAGTCAAGGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACTGAGCCACGTCAA | 480 |
| Db | 3822 | GAAAGTCAAGGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACTGAGCCACGTCAA | 3881 |
| QY | 481 | TGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCAC | 540 |
| Db | 3882 | TGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCAC | 3941 |
| QY | 541 | TCTGGAAGACCTGAAACACCAAGATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAG | 600 |
| Db | 3942 | TCTGGAAGACCTGAAACACCAAGATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAG | 4001 |
| QY | 601 | GCAGCTGCATGAAGCCCAAGGGACCTTTGGTCCAGCATCTCAGCACCTTTCTTCCACGTC | 660 |
| Db | 4002 | GCAGCTGCATGAAGCCCAAGGGACCTTTGGTCCAGCATCTCAGCACCTTTCTTCCACGTC | 4061 |
| QY | 661 | TGTCCAGGTCCTCGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA | 720 |
| Db | 4062 | TGTCCAGGTCCTCGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA | 4121 |
| QY | 721 | CGAGACTCAAACTGCTGGACCCATCCGAGCTTATAGGACTGCCATGAACTCCGAAGACTGCA | 780 |
| Db | 4122 | CGAGACTCAAACTGCTGGACCCATCCGAGCTTATAGGACTGCCATGAACTCCGAAGACTGCA | 4181 |
| QY | 781 | TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCA | 840 |
| Db | 4182 | TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCA | 4241 |
| QY | 841 | GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCCTTGGACCAGCA | 900 |
| Db | 4242 | GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCCTTGGACCAGCA | 4301 |
| QY | 901 | CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC | 960 |
| Db | 4302 | CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC | 4361 |
| QY | 961 | TATTTATGACCGCTGGAGCAAGAGCAACAATTGTTCAACGTCCCTCTCTGCGTGGGA | 1020 |
| Db | 4362 | TATTTATGACCGCTGGAGCAAGAGCAACAATTGTTCAACGTCCCTCTCTGCGTGGGA | 4421 |
| QY | 1021 | TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGT | 1080 |
| Db | 4422 | TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGT | 4481 |

| QY | 1081 | CCTGTCTTTTAAACTGGGCATCATTTCCCTGTGTAAGACACATTTTGAAGACAAGTACAG | 1144 |
|------------|----------------------------|--|---------------------------------|
| Db | 4482 | CCTGTCTTTTAAACTGGGCATCATTTCCCTGTGTAAGACACATTTTGAAGACAAGTACAG | 4541 |
| QY | 1141 | ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT | 1200 |
| Db | 4542 | ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT | 4601 |
| QY | 1201 | CCTTCTGCATGATTCTATCCAAATTCGAAGACAGTTGGTGAAGTTGCATCCTTTGGGG | 1260 |
| Db | 4602 | CCTTCTGCATGATTCTATCCAAATTCGAAGACAGTTGGTGAAGTTGCATCCTTTGGGG | 4661 |
| QY | 1261 | CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTG | 1301 |
| Db | 4662 | CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTG | 4702 |
| RESULT 5 | | | |
| AX538622 | AX538622 | Sequence 42 from Patent WO0229056. | |
| DEFINITION | AX538622 | | |
| ACCESSION | AX538622.1 | GI:25271171 | |
| VERSION | | | |
| KEYWORDS | | | |
| SOURCE | | synthetic construct | |
| ORGANISM | | synthetic construct | |
| REFERENCE | | artificial sequences. | |
| AUTHORS | | 1 | |
| TITLE | | Chamberlain, J.S. and Harper, S.Q. | |
| JOURNAL | | Mini-dystrophin nucleic acid and peptide sequences | |
| | | Patent: WO 0229056-A 42 11-APR-2002; | |
| | | THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US) | |
| FEATURES | | Location/Qualifiers | |
| source | | 1. .8689 | |
| | | /organism="synthetic construct" | |
| | | /mol_type="genomic DNA" | |
| | | /db_xref="taxon:32630" | |
| | | /note="Synthetic" | |
| BASE COUNT | 2721 a | 1804 c | 1861 g 2303 t |
| ORIGIN | | | |
| | Query Match | 100.0%; | Score 1301; DB 6; Length 8689; |
| | Best Local Similarity | 100.0%; | Pred. No. 0; |
| | Matches 1301; Conservative | 0; | Mismatches 0; Indels 0; Gaps 0; |
| QY | 1 | CGACTTTCCAGCAGTTTCAAGACGAGACGATGTATAGGGCCTTCAAGAGGGAATTGAA | 60 |
| Db | 3448 | CGACTTTCCAGCAGTTTCAAGACGAGACGATGTATAGGGCCTTCAAGAGGGAATTGAA | 3507 |
| QY | 61 | AACATAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA | 120 |
| Db | 3508 | AACATAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA | 3567 |
| QY | 121 | GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCTCCTCAGGAGAG | 180 |
| Db | 3568 | GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCTCCTCAGGAGAG | 3627 |
| QY | 181 | AGCCCAGAAATGTACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATCTCAGTGGGA | 240 |
| Db | 3628 | AGCCCAGAAATGTACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATCTCAGTGGGA | 3687 |
| QY | 241 | AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTCAAAGACT | 300 |
| Db | 3688 | AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTCAAAGACT | 3747 |
| QY | 301 | CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT | 360 |
| Db | 3748 | CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT | 3807 |
| QY | 361 | CAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAAAGATCACTCGA | 420 |
| Db | 3808 | CAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAAAGATCACTCGA | 3867 |

QY 181 AGCCAGAAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGAGGTCATAACTAGTGGGA 240
DB 8896 AGCCAGAAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGAGGTCATAACTAGTGGGA 8955

QY 241 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 300
DB 8956 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 9015

QY 301 CCAGGAACCTCAAGAGGCCACGGATCAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 360
DB 9016 CCAGGAACCTCAAGAGGCCACGGATCAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 9075

QY 361 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 420
DB 9076 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 9135

QY 421 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAAAGTGAAGCCACGTCAA 480
DB 9136 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAAAGTGAAGCCACGTCAA 9195

QY 481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 540
DB 9196 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 9255

QY 541 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGCGCGTCGAGGACCGAGTCAG 600
DB 9256 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGCGCGTCGAGGACCGAGTCAG 9315

QY 601 GCAGCTGCATGAAGCCCCACAGGGACCTTTGGTCAGCATCTCAGCACCTTCTTTCCACGTC 660
DB 9316 GCAGCTGCATGAAGCCCCACAGGGACCTTTGGTCAGCATCTCAGCACCTTCTTTCCACGTC 9375

QY 661 TGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA 720
DB 9376 TGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA 9435

QY 721 CGAGACTCAACAAACTTGCTGGGACCAATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
DB 9436 CGAGACTCAACAAACTTGCTGGGACCAATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 9495

QY 781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
DB 9496 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 9555

QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCAATGTGATGCCTTGGACCAAGCA 900
DB 9556 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCAATGTGATGCCTTGGACCAAGCA 9615

QY 901 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAAC 960
DB 9616 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAAC 9675

QY 961 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGCTCCCTCTCTGCGTGGGA 1020
DB 9676 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGCTCCCTCTCTGCGTGGGA 9735

QY 1021 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGCTGT 1080
DB 9736 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGCTGT 9795

QY 1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAGTACAG 1140
DB 9796 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAGTACAG 9855

QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGGCT 1200
DB 9856 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGGCT 9915

QY 1201 CCTTCTGCATGATCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1260
DB 9916 CCTTCTGCATGATCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 9975

QY 1261 CAGTAACATTGAGCCAAAGTGTCCGAGCTGCTTCCAAATTG 1301

Db 9976 CAGTAACATTGAGCCAAAGTGTCCGAGCTGCTTCCAAATTG 10016

RESULT 8
HSDMDR Human mRNA for dystrophin.
LOCUS 12446 bp mRNA linear PRI 12-SEP-1993
DEFINITION
ACCESSION X14298
VERSION X14298.1 GI:30845
KEYWORDS Dmd gene; Duchenne muscular dystrophy; dystrophin.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 12446)
Rosenthal,A., Speer,A., Billwitz,H., Cross,G.S., Forrest,S.M. and
Davies,K.E.
Two human cDNA molecules coding for the Duchenne muscular dystrophy
(DMD) locus are highly homologous
Nucleic Acids Res. 17 (13), 5391 (1989)
89345106
2668885
2 (bases 1 to 12446)
Rosenthal,A.
Direct Submission
Submitted (09-FEB-1989) Rosenthal A., Akademie der Wissenschaften
der DDR, Zentralinstitut fuer Molekularbiologie, Robert-Roessle
Str.10, 1115 Berlin Buch, DDR
see also M18533 and M20250 for Dmd seqs.; discrepancies compared to
M18533 cDNA were located at x14298 pos. 496, 1772, 1965, 2449,
3687, 4229, 4504, 5075, 5332, 5630 and 7194.
Location/Qualifiers
1. .12446
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X chromosomal, Xp21."
/tissue_type="muscle"
/dev_stage="fetal and adult."
99. .11156
/note="dystrophin (AA 1 - 3685)"
/codon_start=1
/protein_id="CAA32479.1"
/db_xref="GI:30846"
/translation="MLWEEVEDCYEREDVQKTFTKWVNAQFSKFGKHENLFSDL
QDGRLLDLEGLTQKLPKEKSTRVHALNVNKNALRVLQNNVLDLVNIGSDIVDG
NHKLTGLIWNIIILHWQVKNMKNIMAGLQPTNSEKILLSWVRQTRNYPQVNVNFT
TSWSDGLALNALIHSHRPDLFDWNSVVCQSQATOLEHAFNIARYQLGIEKLDPEDV
DTYPDKKSILMYITSLFQVLPQQVSEIAEQIEVEMLPPPKVTKEEHFQLHQHYSQ
QITVSLAQGYERTSSPKPRFSYAYTAAVYTTSDPTRSPFSPHLEAPEDKSPGSSL
MESEVNLDRYQTALIEVLSLAEADTLQAQGEISNDVEVVKDQFHTHEGYNMDLTAH
QGRVGNILQLGSKLIGTCLSEDEETEVEQEMNLLNSRWECLRVASMEKQSNLHRVLM
DLQNKLKLNDWLTKEERTKMEEEPLGPDLELKRQVQKHVLOEDLEQQRVN
SLTHMVVVDSESGDHATAALEEQKLVLDGRWANCRTWEDRWVLLQDILLKWRQLTE
EOCLFSAWLSEKEDAVNKIHTTGFKDQNEMLSSLOKLAVLKADLEKKQSMGKLYSIK
QDLLSTLKNKSVTKTEAWLDNFARCDNLVQKLEKSTAQISQAVTTTQPSLTQTVM
ETVTTVTREQILVKHAQBELPPPPPKKQKITVDSEIRKRLDVIDTELHSWITRSEA
VLOSPEPAIFRKEGNSDLKEKNAIEREKAERKRLQDASRSGQALVEQMVEGVNA
DSIKQASEQLNSRIEFCOLLSERLNWLEYQNNIIAFYNLQQLQLEQMTTAEENWLKIQ
PTTPEPTAIKSQLKICKDEVNRLSGLQPQIERUKIQSIALKEKGQGMFLDADFVAF
TNHFQVFSVDQAREKELQTIQPSLNSVNEGGQIKNEAEPPFASRLTELMKELNTQWDMCQVY
EIMEQRLGELQALQSSLOEQSSGLYYLSTTVKEMSKKAPSEISRKYQSEFEIEGRWK
KLSSQLVEHCQKLEEQMNKLRIQNHIIQTLKKWMAEVDVFLKEEWPALGDSEILKKQL
KOCRLLVSDIQTIQPSLNSVNEGGQIKNEAEPPFASRLTELMKELNTQWDMCQVY
ARKEALKGGLEKTVSLQKDLSEHWMTOAEEVEYLRDDEYKTPDELOKAFEMKRAK
EEAQOKEAKVKLLTESVNSVIAQAPPVAQEALEKKELETLTNNYQWLCTRLNGKCKTLE
EYWACWHELLSYLEKANKWLNVEFEKLTNTENIPGAEIIESEVLDSELENMRHSDNP
NQIRILAQTLTDGVMDELINEELETNFSRWREHLEAVRRQKLEQSIQSAQETENS
LHLIQESLTFIDKQLAAYIADKVDAQMPQEAQKIQSDLTSHIISLEEMKKHQGKEA
AQRVLSQIDVAQKQLQDVSMKFRLLFOKPFANFEQRLSKMILDEVKMLPALETKSVE
QEVVQSQNLNHCVNLKSLSEVKSEVEMVIKTGRQIVQKQKTENPKELDERVTALKLHY

FEATURES
source

CDS

NELGAKVTERKQOLEKCLKLSRKMREMNVLTEWLAATDMELTKRSASVEGMPSNLDSE
VAMGKATQETIEKQKVHLKSIITEVGEALKTVLGKKETLVEDKLSLNSNWIATVSRAE
EWLNLILEYQKHMETFDQNVVDHITKWIIOADTLTLLDESEKKPOKEDVLKRLKAE
IRPKVDSTDRQANLMAHGDHCRKLVEPQISELNRFAAISHRIKTKASIPPLKELE
QFNSDIOKLEPLEAEIQQGVNLKEEDFNKMDNEDNEGTVKELLQRGDNLQOORTDER
KSEIEIKIQQLQTKHNLKDLRSQRKKALEISHQWYQYKQOADDLLKCLDDIEKLU
ASLPEPRDERKIEIDRELQKKKEELNVRRAEQLESDGAAAMAVEPTQIOLSKRWE
IESKFAQFRRLNFAQIHTVREETMVMTEDEMPLEISYVPSTVLTEITHVSQALLEVEQ
LNAPDLCAKDFEDLFQEESELKNIKDSIQOSSGRIDIHSHKTAALQSATPVERVKL
QEALSQDLQFQWEKVNMYKDRQGRFDRSEKWRFRHYDIKIFNOMLTEAEQFLRKTQI
PENWEHAKYKWLKELQDQIGQRTVVRTLNATGEEIIQOSSKTDASILEKLGSLNL
RWQEVCKQLSDRKRLKEEQNILEFQRLNEFVFLWLEEDNLTASIPLEPGKEQQLKE
KLEQVLLVEELPLQGIKQLNETGGPVLVSAPISPEEQDKLENKLNKQTNLQWIKVS
RALPEKQGEIEAQIKDLGQLEKLEDELEQLNHLHLLWLSPIRNLQLEIYNQPNQEGPFD
VKETEIAVOAKQPDVEEILSKQHLYKKEPATQPVKRLKLELDSSEWKAVERNLLQELRA
KQPD LAPGLTTIGASPTQTVTLVTPVVTKEBTASIKLEMPSSLMLEVPALADFNRAWT
ELTDWLSLLDQVIKQSRVMVGDLEDINEMI KQKATMODLEQRPQLEELITAAQNKL
NKTNSOEARTIITDRIERIQNOWDEVQEHQNRQQLNEMLKDSQWLEAKEEAEQVL
GOARAKLESWEKGPYVDAIQKKTETKQLAKDLRQWQTVNDVANDLALKLLRDYSAD
NVLQDATTRKLERLLEDSKGVKELMKQWQDLQGEIEAHTDVYHNLNDSQKILRSLEGSD
DAVLLQRRLDNMNFWSLRLKSLNIRSHLEASSDQWKLRLHLSQELLVWLQKODEL
SRQAPIGGDFPAVQKQNDVHRAFKRELKTEPVMSTLETVRIFLTEQPLEGLEKLYQ
EPRELPEERAQNVTRLRLRQKAEVNTWEKLNHLSADWQKIDETLERLQELQEAED
BLDLKLRQAEVIKGSQWPGVGLLIDSLQDHLKVKALRGEIAPLENVSHVNDLARQL
TTLGIQLSPYNLSTLEDNTRWKLQVAVEDVRQLHEAHRDFGASQHLSTSVQGP
WERAI SPNKVPYIYNHETQTTCDWHPKMTLEYQSLADLNNRFSAYRTAMKLRRLQKA
LCGLDLSLSAAACDALDOHNLKONDOPMDILOINCLTIIYDRLEQEHNLNVNPLCVD
MCLNWLNLVYDTRGRIRVLFSFKTGIISLCAHLEDKRYLFKQVASTGFCQDQRRLL
GILLHDSIQIPROLGEVASFGGSNIEPSVSCFPANNKPEIEAALFLDMWRLEPQSM
VWLPVLHRAAAETAHQAKNCNICEPIIGFIRYSLKHFNFDIQCSCFFSGRVAKGH
KMHYPMVEYCTPTTSGEDVRDFAKVLKNKFRKRYFAKHPRMGYLBPVQTVLEGNMET
PVTLINFVPVDSAPASSPOLSHDDTHSRIEYHVASLAEMENSGVINDSISPNEID
DEHLLIQHYCQSLNQDSPLSQSPSPAQILISLSEERELERILADLEENRNLOAEY
DRLKQHEHKGKLSPLPSPPEMPTSPQSPRDAELIAEAKLLRQHKRLEARMQILEDH
NKQLESQHLRLQLLEQPAEAKVNGTTVSSPSTLQSDSSQPMLLRVLVGSQTSDSM
GBEDLLSPQDSTSTGLEEVMQNLNNSFPSSRGRTPGKPMREDTM"

BASE COUNT 4135 a 2524 c 2876 g 2911 t

ORIGIN

Query Match 100.0%; Score 1301; DB 9; Length 12446;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTTCCAGAGTTTCAAGAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAAATTGAA 60
DB 8606 CGACTTTCCAGAGTTTCAAGAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAAATTGAA 8665

QY 61 AACTAAAGAACTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
DB 8666 AACTAAAGAACTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 8725

QY 121 GCCTTTGGAAGGACTAGAGAAACTCTACAGAGAGCCAGAGAGCTGCCTCCTGAGGAGAG 180
DB 8726 GCCTTTGGAAGGACTAGAGAAACTCTACAGAGAGCCAGAGAGCTGCCTCCTGAGGAGAG 8785

QY 181 AGCCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 240
DB 8786 AGCCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 8845

QY 241 AAAATTGAACCTTGCCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 300
DB 8846 AAAATTGAACCTTGCCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 8905

QY 301 CCAGGAACTTCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 360
DB 8906 CCAGGAACTTCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 8965

QY 361 CAAGGGATCTGGGAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCCAAGATCACCTCGA 420
DB 8966 CAAGGGATCTGGGAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCCAAGATCACCTCGA 9025

QY 421 GAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTGAGCCACGTCAA 480

Db 9026 GAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTGAGCCACGTCAA 9085

QY 481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAAACCTCAGCAC 540

Db 9086 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAAACCTCAGCAC 9145

QY 541 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGGAGTCAG 600

Db 9146 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGGAGTCAG 9205

QY 601 GCAGCTGCATGAAGCCCCACAGGGACTTTTGGTCCAGCATCTCAGCACATTTCTTTCCACGTC 660

Db 9206 GCAGCTGCATGAAGCCCCACAGGGACTTTTGGTCCAGCATCTCAGCACATTTCTTTCCACGTC 9265

QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720

Db 9266 TGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 9325

QY 721 CGAGACTCAAAACAACTTGTGTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780

Db 9326 CGAGACTCAAAACAACTTGTGTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 9385

QY 781 TGACCTGAATAATGTCAAGTTCTCAGCTTATAGGACTGCCATGAAACTCCGAGACTGCA 840

Db 9386 TGACCTGAATAATGTCAAGTTCTCAGCTTATAGGACTGCCATGAAACTCCGAGACTGCA 9445

QY 841 GAAGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCCCTTGACCCAGCA 900

Db 9446 GAAGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCCCTTGACCCAGCA 9505

QY 901 CAACCTCAAGCAAAATGACAGCCCATGGATATCTTCAGATATTTAAATTGTTGACCAC 960

Db 9506 CAACCTCAAGCAAAATGACAGCCCATGGATATCTTCAGATATTTAAATTGTTGACCAC 9565

QY 961 TATTATGACCGCTGGAGCAAGAGCACAACAATTTGGTCAACCGTCCCTCTCTCGCTGGA 1020

Db 9566 TATTATGACCGCTGGAGCAAGAGCACAACAATTTGGTCAACCGTCCCTCTCTCGCTGGA 9625

QY 1021 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1080

Db 9626 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 9685

QY 1081 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140

Db 9686 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 9745

QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGCGAGGTGGGCT 1200

Db 9746 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGCGAGGTGGGCT 9805

QY 1201 CCTTCTGCATGATTCTATCCAAATTTCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1260

Db 9806 CCTTCTGCATGATTCTATCCAAATTTCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 9865

QY 1261 CAGTAACATTGAGCCCAAGTGTCCGAGCTGCTTCCAATTG 1301

Db 9866 CAGTAACATTGAGCCCAAGTGTCCGAGCTGCTTCCAATTG 9906

RESULT 9
AX409637
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AX409637
Sequence 2284 from Patent WO0229103.
AX409637
AX409637.1 GI:21442342
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.

AX409637
13957 bp
DNA
linear
PAT 14-JUN-2002

TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2284 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
source Location/Qualifiers
1. .13957
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/notes="EMBL/GenBank Accession No. M18533"
BASE COUNT 4602 a 2781 c 3122 g 3452 t
ORIGIN
Query Match 100.0%; Score 1301; DB 6; Length 13957;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCCAGCAGTTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
Db 8716 CGACTTTCCAGCAGTTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 8775
QY 61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db 8776 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 8835
QY 121 GCCTTTGGAAGGACTAGAGAACTCTACCAGGAGCCAGAGAGTGCCTCCTTGAGGAGAG 180
Db 8836 GCCTTTGGAAGGACTAGAGAACTCTACCAGGAGCCAGAGAGTGCCTCCTTGAGGAGAG 8895
QY 181 AGCCCGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 240
Db 8896 AGCCCGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 8955
QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 300
Db 8956 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 9015
QY 301 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT 360
Db 9016 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT 9075
QY 361 CAAGGGATCCTGGCAGCCCGTGGGCGATCTCTCATTTGACTCTCTCCAAGATCACCTCGA 420
Db 9076 CAAGGGATCCTGGCAGCCCGTGGGCGATCTCTCATTTGACTCTCTCCAAGATCACCTCGA 9135
QY 421 GAAAGTCAAGGCACCTTCAGGAGAGAAATTGCGCCTCTGTAAAGAGAACGTGAGCCACGTCAA 480
Db 9136 GAAAGTCAAGGCACCTTCAGGAGAGAAATTGCGCCTCTGTAAAGAGAACGTGAGCCACGTCAA 9195
QY 481 TGACCTTGCTGGCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCAC 540
Db 9196 TGACCTTGCTGGCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCAC 9255
QY 541 TCTGGAAGACCTGGAACACCCAGATGGAAGCTTCTGCAGGTGGCGCTCGAGGACCGAGTCAAG 600
Db 9256 TCTGGAAGACCTGGAACACCCAGATGGAAGCTTCTGCAGGTGGCGCTCGAGGACCGAGTCAAG 9315
QY 601 GCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
Db 9316 GCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 9375
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
Db 9376 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 9435
QY 721 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
Db 9436 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 9495
QY 781 TGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
Db 9496 TGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 9555
QY 841 GAAGGCCCTTTGCTTGGATCTCTTTGAGCCCTGTGAGCTGCATGTGATGCCCTTGGACCAGCA 900

Db 9556 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCTTGGACCAGCA 9615
QY 901 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCAC 960
Db 9616 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCAC 9675
QY 961 TATTTATGACCGCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCTCTCTGCGTGA 1020
Db 9676 TATTTATGACCGCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCTCTCTGCGTGA 9735
QY 1021 TATGTCTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1080
Db 9736 TATGTCTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 9795
QY 1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
Db 9796 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 9855
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAAACAGGATTTTGTGACCGCAGGCTGGGCCT 1200
Db 9856 ATACCTTTTCAAGCAAGTGGCAAGTTCAAACAGGATTTTGTGACCGCAGGCTGGGCCT 9915
QY 1201 CCTTCTGCATCATTTCTATCCAAATTCGAAGACAGTTGGTGAAGTTGCATCCTTTGGGG 1260
Db 9916 CCTTCTGCATCATTTCTATCCAAATTCGAAGACAGTTGGTGAAGTTGCATCCTTTGGGG 9975
QY 1261 CAGTAACATTGAGCCCAAGTGTCCGAGCTGCTTCCAATTG 1301
Db 9976 CAGTAACATTGAGCCCAAGTGTCCGAGCTGCTTCCAATTG 10016
RESULT 10
AX538581
LOCUS AX538581 13957 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 1 from Patent WO0229056.
ACCESSION AX538581
VERSION AX538581.1 GI:25271086
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Chamberlain,J.S. and Harper,S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 1 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
source Location/Qualifiers
1. .13957
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 4602 a 2781 c 3122 g 3452 t
ORIGIN
Query Match 100.0%; Score 1301; DB 6; Length 13957;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCCAGCAGTTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
Db 8716 CGACTTTCCAGCAGTTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 8775
QY 61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db 8776 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 8835
QY 121 GCCTTTGGAAGGACTAGAGAACTCTACCAGGAGCCAGAGAGTGCCTCCTTGAGGAGAG 180
Db 8836 GCCTTTGGAAGGACTAGAGAACTCTACCAGGAGCCAGAGAGTGCCTCCTTGAGGAGAG 8895
QY 181 AGCCCGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 240
Db 8896 AGCCCGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 240

Db 8896 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 8955

QY 241 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTTGAAAGACT 300

Db 8956 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTTGAAAGACT 9015

QY 301 CCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 360

Db 9016 CCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 9075

QY 361 CAAGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 420

Db 9076 CAAGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 9135

QY 421 GAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACGTTGAGCCACGTCAA 480

Db 9136 GAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACGTTGAGCCACGTCAA 9195

QY 481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCAC 540

Db 9196 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCAC 9255

QY 541 TCTGGAAGACCTGAAACACCAGATGGAGCTTCTGCAGGTGGCGTTCGAGGACCGAGTCAG 600

Db 9256 TCTGGAAGACCTGAAACACCAGATGGAGCTTCTGCAGGTGGCGTTCGAGGACCGAGTCAG 9315

QY 601 GCAGCTGCATGAAGCCACACAGGGACTTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGTTC 660

Db 9316 GCAGCTGCATGAAGCCACACAGGGACTTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGTTC 9375

QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720

Db 9376 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 9435

QY 721 CGAGACTCAAACAACTTGCTGGGACCATCCCAAATGACAGAGCTCTACCACTCTTTAGC 780

Db 9436 CGAGACTCAAACAACTTGCTGGGACCATCCCAAATGACAGAGCTCTACCACTCTTTAGC 9495

QY 781 TGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840

Db 9496 TGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 9555

QY 841 GAAGGCCCTTTGCTGGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCTTGGACCAGCA 900

Db 9556 GAAGGCCCTTTGCTGGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCTTGGACCAGCA 9615

QY 901 CAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTTGTTGACCAC 960

Db 9616 CAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTTGTTGACCAC 9675

QY 961 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGGA 1020

Db 9676 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGGA 9735

QY 1021 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAGGAGGATCCGCTG 1080

Db 9736 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAGGAGGATCCGCTG 9795

QY 1081 CCTGTCTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140

Db 9796 CCTGTCTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 9855

QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCGCAGGCTGGGCCT 1200

Db 9856 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCGCAGGCTGGGCCT 9915

QY 1201 CCTTCTGCATGATTTCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1260

Db 9916 CCTTCTGCATGATTTCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 9975

QY 1261 CAGTAACATTGAGCCCAAGGTCCGGAGCTGCTTCCAATTTG 1301

Db 9976 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTTG 10016

RESULT 11

HUMDYS

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

FEATURES

source

gene

CDS

13957 bp mRNA linear PRI 25-MAY-2000

Homo sapiens dystrophin (DMD) mRNA, complete cds.

M18533 M17154 M18026 M20250

M18533.1 GI:181856

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1699)

Koenig, M., Hoffman, E.P., Bertelson, C.J., Monaco, A.P., Feener, C. and Kunkel, L.M.

Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals

Cell 50 (3), 509-517 (1987)

87273512

3607877

2 (bases 1678 to 3830)

Hoffman, E.P., Monaco, A.P., Feener, C.C. and Kunkel, L.M.

Conservation of the Duchenne muscular dystrophy gene in mice and humans

Science 238 (4825), 347-350 (1987)

88018015

3659917

3 (bases 1 to 13957)

Koenig, M., Monaco, A.P. and Kunkel, L.M.

The complete sequence of dystrophin predicts a rod-shaped cytoskeletal protein

Cell 53 (2), 219-226 (1988)

88194521

3282674

On May 25, 2000 this sequence version replaced gi:340693.

Draft entry and computer-readable sequence kindly provided by M.Koenig, 01-APR-1988 The severity of muscular dystrophy is determined by the size of the deleted DNA segment. Deletions found in different patients were from positions 302-2200, 473-1168, 1691-1810, and 1169-3011.

Location/Qualifiers

1. 13957

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="X"

/map="Xp21.3-p21.1"

/tissue_type="muscle"

/dev_stage="fetus"

/note="G00-119-850"

1. 13957

/gene="DMD"

209. 11266

/gene="DMD"

/codon_start=1

/product="dystrophin"

/protein_id="AAA53189.1"

/db_xref="GI:181857"

/translation="MLWEEVEDCYEREDVQKKTFTKWVNAQFSKFGKQHIENLFSDL ODGRLLDLLEGLTGOKLPKEGSTRVHALNNVNKALRVLQNNNVDLVNIGSTDIVDG NHKLTGLIWNIIHWOVKNVKNIMAGLQQTNSEKILLSWVRQSTRNYPQVNVINFT TSWSDGLAINALIHSHRPDLFDWNSVCCQSQATQRLEHAFNIARYQLGIEKLLDPEDV DTTYDPKKSILMYITSLFVLPOQVSIIEAIOEVEMLPKPKVTKEEHFQLHHQMHYSQ OITVSLAQGYERTSSPKPRFKSYATQAAVYVTTSDPTSPSPSQHLEAPEDKSGSSL MESEVNDRYQTALAEVLWLLSAEDTLQAQGEISNDVEVVKDQFHTHEGYMDLTAH OGRVGNILQLGSKLIGTKLSEDETEVEQNMNLLNSRWECLRVASMEKQSNLHRLVM DLQNKIKELNDWLTKTEERTKMEEEPLGPDLEDLKRQVQKHVLOEDLEQEQVRVN SLTHMVVVDSESGDHATAALEEQKLVIGDRWANI CRWTEDRWVLLQDILLKWRLTE EQCLFSAWLSEKEDAVNKIHTGTGFKDQNEMLSSQLKLA VLKADLEKKQSMGLYSLK QDILLSTLKNKSVTQKTEAWLDNFARCDNLVQKLEKSTAQISQAVTTTQPSLTQTVM

| | | | | | | | |
|---|------|--|--------------|------------|---------------|-------------------|--------|
| EVTVTVTREILVKHAQEELPPPPQKKRQITVDSSEIRKRLDVIDITELHSHWITRSESE VLQSPFEFAIFRKEGNFSDLKEKVNAIEREKAERFKLQDASRSQAALVEQMVEGVNA DSIKQASEQLNSRWIEFCQLLSERLNMWLEYQNNIIFAYNQNLQOLEQMVTTFIENWLKIQ PTTSEPTAIKSQIKICKDEVNRLSGLQPOIERLKIQSIQALKEKGQPMFLDADFAV TNHPKQFSDVOAREKELQITFDTLPPMRYQETMSAIRTWVQSETKLSIPQLSVTDY EIMEQRLGELQALQSSLOEQSGLYYLSTTVKEMSKAPSEISIRKYQSEFEIEIEGRWK KLSSQLVHERCQKLEEQMNKLRKIQNHIOQLTKKMAEVDVFLKEEPALGDSHILKQL KQCRLLVSDIQTIOPLSLNSVNEGQIKNEAPEFASRLTELETKELNTQWDMHCQQVY ARKEALKGGLSTEQKLDJSEMHEWMTQAEFEYELTDEYKTPDELOKAVBEMKRAK BEAQKEAKVKLLLTESVNSVIAQAPPVAQEAALKKELETLTNTYQWLCTRLNGKCKTLE EWACWHELLSYLEKANKWLNVEVFKLTKTENIPGAAEISEVLDLSLNSMRHSEDNP NQIRILAQTLTDGGVMDLINELETFNSRWRELHEEAVRRQKLLLEQSIQSAQETESK LHLIQESLTFIDKQLAAYIADKYDAAQMPQEAQKIOSDLTSHIEISLEEMKKNQKKEA AQVLSQIDVAQKQLQDVSMKFLRFQKPAFELRLQESKMILDEVKHLPALETKSVE QEVVQSLNHCNVLYKSLSEVKSEVEMVKTGRQIQVKQMLTEKSLSAVEGMPNSLDS NELGAKVTERKQOQLEKCLSRKMKEMNVLTWLAATMDMELTKRSAGEVMPNSLDS VAWGKATKEIEKQKVHLKSIIEVGEALKTVLGKETLVEDKLSLNSNWIATVSRAE EWNLLLEYQKHMETFQNVHDHTKWIQADTLDESEKKKPKQEKEDVLRKLKAE IRPKVDSTRDQAANLMANRGDHCGLVPEQISELNHRFAAISHRITKGKASIPKLE QFNDSIQKLEPLEAEIQQVNLKBEEDFNKOMNEDNKGVELLQKQGNLQQRITDEL KREBIKIQKQLLOTKHNALKDLRSQRRKKALEISHQWYQYKROADDLKLCLDIEKKL ASLPEPRDERKIKEIDRELQKKKEELNARRQAEGLSDEGAAVAVEPTQIQLSKRWRE IESKFAQFRRNLNFAQIHTVREETMMVTEDMPLEISYVPSVTYLTETHVQSALLEVEQ LLNAPDLCAKQFEDLFKQEEESLKNIKDSLQSSGRIDIHKSKTAAALQSATPVERVK QEALSQDLQFQWEKVNKMYKDRQGRFDRSVEKWRRFYHDIKIFNQWLTEAEQFLRKTQI PENWEHAKYKWLKELQDQIGORQTVVRLTNATGEEIIQOSSKTDASILEKELGSLNL RWQEVCKQLSDRKKRLEEQNILSEFQDLNEFVLWLEADNIAISIPLEPKEQOOLKE KLEQVKLLVEELPLRQGLKQLNETGGPLVSAPISEEDQKLENKLTQNLQWIKVS RALPEKQGEIEAQIKDLQLEKKLEDEEQNLHLLLWLSPIRNQLKLIYNQNPQEGPFL VQETEIAVQAKQPDVEEILSKQHLKKEKPATQPVKRLLEDLSWKAVNRLLOELRA KQPD LAPGLTTIGASPTQTVTLVQPVVTKETAISKLEMPSSLMLEVPALADFNRAWT ELTDWLSLLDQVIKSORVMVGLEDINEMI IKOKATQDLEQRRPOLEELITAAQNLK NKTNSQEA RTIIITDRIERIQNWQDEVQEHQLNRRQQLNEMLKDSTQWLEAKEEAEQVL GQARAKLESWKGPYTVDAIQKKITETQLAKDLRQWQTNVDVANDLALKLLRDYSAD DTRKVHMITENINASWRSIHKRVSEBEEALAEETHRLLQFPDLLEKFLAWLTEAETTA NVLQDATRKERLLEDSSKGVKELMKQWQDLQGEIEAHTDVYHNLNDSQKILRSLEGS DAVLLQRRLDNMNFKWSLRKXSLNIRSHLEASSDQWKRLHLSLQELLVWLQKDEL SROAPIGGDFFPAVQKQNDVHRAFKRELTKPEVIMSTLETVRIIFLTPLEGLEKLEYQ EPRELPPPEERAQNVTRLRKQAEVNTWEKLNHLSADWQRKIDETLERLOELQEA ELDLKLRAEVIKGSWQPVGDLILSLQDHLKVKALRGEIAPUKENVSHVNDLARQL TILGILQSLPYNLSTLEDLNRWLKLQVAVEDRVRLHEAHRDFGASQHFAMTLRLQKA WERAI SPNKVPYVINHETQTTCTWDHPKMTELYQSADLNNVRFSAYETAKMLRLQKA LCLDLLSLSAACDALDOHNLKQNDQPMIDLIINCLTIIYDRLEQHNHNVNPLCVD MCLNWLNVYDTRGTRIRVLSFKTGIISLCKAHLEDKYRFLFKVASSTGFCQDQRRLL GILLHDSIQIPRQLGEVASFGSNIPEPSVRS CFQFANNKPEIEAALFDLWMLRLEPQSM WVLPVHRVAAAEATAKHQAKNICKECPIIGFGRYSLKHFNYDIIQSCFFFSGRVAKGH KMHPVMVEYCTPTTSGEDVDRDFAKVLNKNKFRTRYFAKHPRMGYLPVQTVLEQDNMET PVTILNFWPVD SAPASSPQLSHDDTHSRIEHYASRLAEMENSGSYLNDISIPNESID DEHLIIQHYCOSLNQDSPLSQRSPAQILISLESEERGELERILADLEENRNLCQAEY DRLKQOQHEKGLSPLSPPEMPTSPQSPPRDAELIAEAKLLRQHKRLEARMQIILEDH NKQLESQHLRLRQLLEQPOAEAKVNGTIVTSSPSTLSRSDSSQPMLLRVVVSQTSDSM GEEDLLSPPDQDTSTGLEEVMEQLNNSFPSSSRGRNTPGKPMREDT M" | | BASE COUNT | 4602 a | 2781 c | 3122 g | 3452 t | ORIGIN |
| Query Match | | 100.0%; | Score 1301; | DB 9; | Length 13957; | | |
| Best Local Similarity | | 100.0%; | Pred. No. 0; | | | | |
| Matches 1301; | | Conservative | 0; | Mismatches | 0; | Indels 0; Gaps 0; | |
| QY | 1 | CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA | 60 | | | | |
| Db | 8716 | CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA | 8775 | | | | |
| QY | 61 | AACTAAAGAACCTGTGAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA | 120 | | | | |
| Db | 8776 | AACTAAAGAACCTGTGAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA | 8835 | | | | |
| QY | 121 | GCCTTTGGAAGGACTAGAGAAACTCTACCAAGGAGCCACAGAGAGCTGCCTCCTGAGGAGAG | 180 | | | | |
| Db | 8836 | GCCTTTGGAAGGACTAGAGAAACTCTACCAAGGAGCCACAGAGAGCTGCCTCCTGAGGAGAG | 8895 | | | | |
| QY | 181 | AGCCACGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA | 240 | | | | |

| | | | |
|----|------|---|------|
| Db | 8896 | AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA | 8955 |
| QY | 241 | AAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTTGAAAAGACT | 300 |
| Db | 8956 | AAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTTGAAAAGACT | 9015 |
| QY | 301 | CCAGGAACTTCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT | 360 |
| Db | 9016 | CCAGGAACTTCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT | 9075 |
| QY | 361 | CAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCTCCAAGATCACCTCGA | 420 |
| Db | 9076 | CAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCTCCAAGATCACCTCGA | 9135 |
| QY | 421 | GAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAAGAGAACGTGAGCCACGTCAA | 480 |
| Db | 9136 | GAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAAGAGAACGTGAGCCACGTCAA | 9195 |
| QY | 481 | TGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC | 540 |
| Db | 9196 | TGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC | 9255 |
| QY | 541 | TCTGGAAGACCTTGAACACCAAGATGGAAGCTTCTGCAGGTGGCGCTGAGGACCGGAGTCAG | 600 |
| Db | 9256 | TCTGGAAGACCTTGAACACCAAGATGGAAGCTTCTGCAGGTGGCGCTGAGGACCGGAGTCAG | 9315 |
| QY | 601 | GCAGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGACACTTTCTTTCCACGTC | 660 |
| Db | 9316 | GCAGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGACACTTTCTTTCCACGTC | 9375 |
| QY | 661 | TGTCCAGGTCCTTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA | 720 |
| Db | 9376 | TGTCCAGGTCCTTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA | 9435 |
| QY | 721 | CGAGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC | 780 |
| Db | 9436 | CGAGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC | 9495 |
| QY | 781 | TGACCTGAATAATGTCTAGATTCTCAGCTTATAGGACTGCCATGATGATGATGATGATGATGAT | 840 |
| Db | 9496 | TGACCTGAATAATGTCTAGATTCTCAGCTTATAGGACTGCCATGATGATGATGATGATGATGAT | 9555 |
| QY | 841 | GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGATGATGATGATGATGATGATGAT | 900 |
| Db | 9556 | GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGATGATGATGATGATGATGATGAT | 9615 |
| QY | 901 | CAACCTCAAGCAAAATGACAGCCCATGAGTATCCTGCAGATTATTAATTTGTTGACCCAC | 960 |
| Db | 9616 | CAACCTCAAGCAAAATGACAGCCCATGAGTATCCTGCAGATTATTAATTTGTTGACCCAC | 9675 |
| QY | 961 | TATTTATGACCGCTGGAGCAAGAGCAACAAATTTGGTCAACGTCCTCTCTGCGTGA | 1020 |
| Db | 9676 | TATTTATGACCGCTGGAGCAAGAGCAACAAATTTGGTCAACGTCCTCTCTGCGTGA | 9735 |
| QY | 1021 | TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGT | 1080 |
| Db | 9736 | TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGT | 9795 |
| QY | 1081 | CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAG | 1140 |
| Db | 9796 | CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAG | 9855 |
| QY | 1141 | ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGGAGGCTGGGCT | 1200 |
| Db | 9856 | ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGGAGGCTGGGCT | 9915 |
| QY | 1201 | CCTTCTGCATGATTTATCCAAATTCGAAGACAGTTGGGTGAAGTTGTCATCCTTTGGGG | 1260 |
| Db | 9916 | CCTTCTGCATGATTTATCCAAATTCGAAGACAGTTGGGTGAAGTTGTCATCCTTTGGGG | 9975 |
| QY | 1261 | CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTTG 1301 | |
| Db | 9976 | CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTTG 10016 | |

RESULT 12
AR220819
LOCUS AR220819 13977 bp DNA linear PAT 26-SEP-2002
DEFINITION Sequence 60 from patent US 6426186.
ACCESSION AR220819
VERSION AR220819.1 GI:23327696
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 13977)
AUTHORS Jones,K.A., Volkmuth,W. and Walker,M.G.
TITLE Bone remodeling genes
JOURNAL Patent: US 6426186-A 60 30-JUL-2002;
FEATURES
Location/Qualifiers
1. .13977
/organism="unknown"
BASE COUNT 4596 a 2765 c 3120 g 3453 t 43 others
ORIGIN
Query Match 99.2%; Score 1290; DB 6; Length 13977;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 CGACTTTCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCCTCAAGAGGGGAATTGAA 60
Db |||||
QY 8716 CGACTTTCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCCTCAAGAGGGGAATTGAA 8775
Db |||||
QY 61 AACTAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA 120
Db |||||
QY 8776 AACTAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA 8835
Db |||||
QY 121 GCCTTTTGAAGGACTAGAGAAACTCTACCAAGAGCCCGAGAGCTGCCTCCTGAGGAGAG 180
Db |||||
QY 8836 GCCTTTTGAAGGACTAGAGAAACTCTACCAAGAGCCCGAGAGCTGCCTCCTGAGGAGAG 8895
Db |||||
QY 181 AGCCCAGAATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATACTGAGTGGA 240
Db |||||
QY 8896 AGCCCAGAATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATACTGAGTGGA 8955
Db |||||
QY 241 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 300
Db |||||
QY 8956 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 9015
Db |||||
QY 301 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT 360
Db |||||
QY 9016 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT 9075
Db |||||
QY 361 CAAGGGATCTCTGGCAGCCCGTGGCGATCTCTCATTTGACTCTCTCAAGATCACCTCGA 420
Db |||||
QY 9076 CAAGGGATCTCTGGCAGCCCGTGGCGATCTCTCATTTGACTCTCTCAAGATCACCTCGA 9135
Db |||||
QY 421 GAAAGTCAAGGCACCTTCAGGAGAGAAATTGGCGCTCTGAAAGAGAACGTGAGCCACGTCAA 480
Db |||||
QY 9136 GAAAGTCAAGGCACCTTCAGGAGAGAAATTGGCGCTCTGAAAGAGAACGTGAGCCACGTCAA 9195
Db |||||
QY 481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 540
Db |||||
QY 9196 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 9255
Db |||||
QY 541 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAG 600
Db |||||
QY 9256 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAG 9315
Db |||||
QY 601 GCAGTGCATGAAGCCCAAGGACCTTTGGTCCAGCATCTCAGCACCTTCTTTCCACGTC 660
Db |||||
QY 9316 GCAGTGCATGAAGCCCAAGGACCTTTGGTCCAGCATCTCAGCACCTTCTTTCCACGTC 9375
Db |||||
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
Db |||||
QY 9376 TGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 9435
Db |||||

QY 721 CGAGACTCAAAACAACTTGCTGGGACCATCCAAAATGACAGAGCTCTACCAAGCTCTTAGC 780
Db |||||
QY 9436 CGAGACTCAAAACAACTTGCTGGGACCATCCAAAATGACAGAGCTCTACCAAGCTCTTAGC 9495
Db |||||
QY 781 TGACCTGAATAATGTGAGATTCTCAGATTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
Db |||||
QY 9496 TGACCTGAATAATGTGAGATTCTCAGATTATAGGACTGCCATGAAACTCCGAAGACTGCA 9555
Db |||||
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGATGATGCCTTGACCCAGCA 900
Db |||||
QY 9556 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGATGATGCCTTGACCCAGCA 9615
Db |||||
QY 901 CAACCTCAAGCAAAAATGACCCCATGGATATCCTGAGATTATTAATTGTTGACCAC 960
Db |||||
QY 9616 CAACCTCAAGCAAAAATGACCCCATGGATATCCTGAGATTATTAATTGTTGACCAC 9675
Db |||||
QY 961 TATTTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGA 1020
Db |||||
QY 9676 TATTTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGA 9735
Db |||||
QY 1021 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 1080
Db |||||
QY 9736 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 9795
Db |||||
QY 1081 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGACAAGTACAG 1140
Db |||||
QY 9796 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGACAAGTACAG 9855
Db |||||
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCT-GGGCC 1199
Db |||||
QY 9856 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGGC 9915
Db |||||
QY 1200 TCCTTCTGCATGATTCTATCCAAATTCGAAGACAGTTGGTGGTGAAGTTGCATCCTTTGGG 1259
Db |||||
QY 9916 TCCTTCTGCATGATTCTATCCAAATTCGAAGACAGTTGGTGGTGAAGTTGCATCCTTTGGG 9975
Db |||||
QY 1260 GCAGTAACATTGAGCCCAAGTGTCCGAGCTGCTTCCAATTG 1301
Db |||||
QY 9976 GCAGTAACATTGAGCCCAAGTGTCCGAGCTGCTTCCAATTG 10017
Db |||||

RESULT 13
AF070485
LOCUS AF070485 13887 bp mRNA linear MAM 09-DEC-1998
DEFINITION Canis familiaris dystrophin mRNA, complete cds.
ACCESSION AF070485
VERSION AF070485.1 GI:3982750
KEYWORDS
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 13887)
AUTHORS Carville,K.S., Mann,C.J., Schatzberg,S.J. and Wilton,S.D.
TITLE Direct Submission
JOURNAL Submitted (04-JUN-1998) ANRI, Pathology, University of Western
Australia, Verdun Street, Nedlands, WA 6018, Australia
FEATURES
Location/Qualifiers
1. .13887
/organism="Canis familiaris"
/mol_type="mRNA"
/strain="breed Golden retriever"
/db_xref="taxon:9615"
5'UTR <1. .289
CDS 290. .11332
/codon_start=1
/product="dystrophin"
/protein_id="AAC83646.1"
/db_xref="GI:3982751"
/translation="MLWEEVEDCYEREDVQKKTFTKWVNAQFSKFGKHLENLFSDL
QDGRRLDLLEGLTGKLPKEKGSRTVHALNNVKNALRQKNNVDLVNIGSTDIVDG
NHKLTLGLIWNIIILHWQKVNVMKINMAGLQQTNSEKILLISWVRQSTRNYPQVNVINFT

TSWSDGLALNALIHSHRPDLFDWNSVVCQSQSATORLEHAFNIAKYQLGIEKLLDPEDV
ATTVPDKSILMYITSLFVLPQQVSI EAIQEVEMLP RPSPQVTR EEHFQIHHQMHYSQ
QITVSLAQYPRFPKFSYAYTQAAVYTTSDPTRSPQSHLETPEDKSFGRS
LTETEANLDSYQATALEEVLWSLLSAEDALQAQGEISNDVEEVKEQFHTHEGYMDLTS
HQGRVGNVLQGSQILGTLSEDEETEVQOMNLNLSRWECLRVASMEKQSNLHKVL
MDLQNOQLKELNDWLTKTEERTRKMEKEPLGPDIEDLKRQVQKHVQLQEDLEQEQVRV
NSLTHMVVVDES GDHAAALEEQLVKVLGDRWANI CRWETDRWLLQDILLKQWRFT
EEQCLFSAWLSEKEDAVNKIHTTGFKDQSEVL SNLQKLA VLKTDLEKKKQTMDKLCSL
NQDLLSALKNTVVAHKMEAWLNDNFAQRWNLVQKLKSSAQISOAVTTQPSLTQTTV
MEITVMVTTREHILVKHAQEEELPPPPQKKRQIIVDSEIRKRLDVIDTELHSWITRSE
AVLQSPFEAIYRKEGNSDLKEKNVAIEREKAEPKRLQDASRSAQALVEQMNEGYN
ADSIKQASEQLNSRWIEFCQLLSERLNLWLEYONNIITFYNLQQLQLEQMTTAEANWLKT
OPTTSEPTAIKSQ LKICKEINRLSALQPQIERIKIQSIALKEKGQGPFLDADFVA
FTNHFNQVFADVQAREKELQITFDLSLPPMYQETNSTILTWIQQSETKLSIPQVTVTE
YDIMEORLQELQALQSSLOEQONGNLNLSSTTVKEMSKAPLSDISRKYQSEFEIEGR
WKLSSQLVEHCQKLEEQMAKLKRIQNHIKTLKKWITEVDVFLKEEWPALGDSEILKR
QLKQCRLLVNDIOTIQPSLNSVNEGAKQMKNEAEPFAGRLTELRELNTQWDMCRQ
VYARKEALKGGLDKTVSLQDLSEMHWMTOAEEEYLERDFEYKTPDELOQTAVEEMKR
AXEEAQCKEAKVKLLLTESVNSVIAQAPPAQAEALKKELDLTTNYQWLCTRLNGCKKT
LEEYVACWHHELLSYLEKANKWLSEVEVKLKTNTENISGGAEEIAEVLDSLENLMQHSIE
NPNOIRILAQTLTLDGVMDELINEELETENSRWRLEHEAVRRQKLLQESIQSAQIE
KSLHLIQESLSSIDKQLAAYIADKVDAQAQMPQEAQIKQSDLTSHEISLEEMKKNQKG
ETAQRVLSQIDVAQKQLQDVSMKFRFLQKPANFEQRLQESKMILDEVKXHLPALETKS
VEQEVVQSNLHCNVNLKSLSEVKSEVEMVIKTQIQIQKKQNTENPKELDERVTALKL
HYNELGAKVTERKQOQLEKCLKLSRKMRKEMNALTEWLAATDMELTKRSAVEGMPNSLD
SEVANGKATQKEIEKQKHLKSVTEVGEALKTVLKGKMELVEDKLSLNSNWI AVT SR
AEEWNLALLEYQKHMETFDQNVYITNWIIQADALLDESEKKKPOQKEDILKRLKAEM
NDIRPKVDANLANRGDHCRCRVVEPKISLNHRFAAISHRITKTKGASIQRLTE
LEQPNSDIOKLLPLEAEIQQGVNLKEEDFNKDMSEDNQGVTKELLQRGDNLQIPRIDE
ERKBEIKIKOOLLQTKHNALKDLRSORRKALETSHQWYQYKQRAODDLKCLDDIEK
KLASLPEPRDERKIKEIDRELQKKEELNAVRRQAEGLSDEGAAAVEPTQIQLSKRW
RETESKPAQFRLNFAQIHTVHEESVWAMTEDMPLEISYVPTSLTELTHSVSQAELSEV
EELNAPDLCAQDFEDLFQEEESLKNIKDSLOQISGRIDI IHNKKTAALHSATPAERA
KLOEALSRLDFQWERVNNMYKDRQGRFDRSVEKWRPFHYDMKILNQLWLTEAEQFLKKT
QIPENWEHAKYKWLKELQDQIGQORQSVRVNLNATGEEIIQSSKTDASILQEKLGSL
NLRWQEVCKQLAERKRLEROKNILSEFORDVNEFLWLWEEADNVANTPLEPNEGQOL
KEKLEBQVKLLAEELPLRQGLKQLNETGGTVLVSAPLSPEEQDKLENKLNQTNLQWIK
VSRNLPEKQEEIEAHVXDLQLEELNHLILLWSLPRNQLEIYINQPNQTPGFDIKEIE
VAVQAKQPDVEGILSKQHLYKEKPATQPAKRLKLEDSSDWKVVTLQLELRAKQPGP
APGLTTVRAPPSQTVTLVTQPAVTKETAISKLEMPSSLLLEVPALADFNRAWTELTDW
LSELDRIKSVQVMVGDELINEMI IKQATLQLEQRRPOLEELITAAQNLKNKTSN
QEARTIITDRIERIQSQWDEVQEHQNRRLQLTMLKDSQWLEAKEEAEQVLQGARA
KLESWKEAPYTVDAIOKKTETKQLAKQLRQWQINVDVANDLALKLLRDYSADDDRKY
HMITENINASWASIHKRLSREAALEETHRLLQOPPLDLEKFLAWLTHEATTANVLQD
ATHXERILLEDSKGVRELKMQDLQGEIEAHTDIYHNLDENGQKVLRSLEGSDDAALL
QRLLDNMNFKWSELKKSINIRSHLEASSDQWKRLHLSLOELLVWLQKDDLSRQAP
IGGDFPAVQKQNDVHRAFKRELKTEPVMISTLETVRIFLTEQPLEGLEKLYQEPREL
PPEERAQNVTRLRLKQAEVNTQWEXKLVNHSADQWKIDAEALERLOELQEATDELDLK
LRQAEVKGSWQPVGDLLISLDQHLKVKALRGITPLKENVSYYVNDLARQLTTLGI
QLSPYNLNTLEDLNRWKLQVAIEDRIROLHEAHRDGPASQHFLSVSVQGPWERAI
SPNKVPYINHETQTCWDHPKMTELYQSLADLNNRWFSAYRTAMKLRRLQKALCLDL
LSLSAACDALDQHNKQNDQPMDILQVINCLTTIYDRLEQEHNNLVNVLCDVMCLNW
LLNVYDTGRGIRVLSRKTGIIISLCAHLEDKYRYLFKQVASSSTGFCQRRRLGLLH
DSIQIPROLGEVASFSGGNIIEPSVRSCFPANNKPEIEAALFLDWMRLPEQSMVWLPV
LHRVAAAEATAKHQAKCNICEKCPIIIGFRYRSLKHPNYDICIQSCFFSGRVAKGHKMHYP
MVEYCTPTTSGEDVRDFAKVLKNKFRTKRYFAKHPRMGYLPVQTVLEGNNMETPVTLI
NFWPVD SAPASSPOLSHDDTHSRIEHYASRLKXMNENSGSYLNDISI SPNESIDDEHLL
IQHYWRSLNQESPLSQPRSPAQILISLESEERGELERILADLEGRNRLQAEYDRLKQ
QHEHKGLSPLPSPPEMPTSPQSPRDAELIAEAKLLRQHKGRLEARMQILEDHKNQLE
SQLHLRLQLEQPOAEAKVNGTTVSSPSTSLQRSDSSQPMLLRRVVVGSQTSSEMGEDL
LSPQDSTSTGLEEVMEQLNHSFPSSRGRNTPGKPNREDTM"

3'UTR 11333. >13887
repeat_region 11789. .11828
/rpt_type=tandem
/rpt_unit=ca
BASE COUNT 4455 a 2906 c 3164 g 3362 t
ORIGIN

Query Match 88.7%; Score 1154.4; DB 4; Length 13887;
Best Local Similarity 93.0%; Pred. No. 0;
Matches 1209; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 2 GACTTTCAGCAGTTCAGAAAGCAGACGATGTACATAGGGCCTTCAAGAGGGAATTGAAA 61

| | | | |
|----|------|--|------|
| Db | 8783 | GACTTTCAGCGGTGCAGAAAGCAGAAATGATGTACACAGGGCCTTCAAGAGGGAATTGAAA | 8842 |
| QY | 62 | ACTAAAGAACTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAG | 121 |
| Db | 8843 | ACGAAAGAACTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAG | 8902 |
| QY | 122 | CCTTTGGAAGGACTAGAGAAACTCTACAGGAGCCCGAGAGAGTGCCTCCTGAGGAGAGA | 181 |
| Db | 8903 | CCTTTAGAGGACTAGAGAAACTCTACAGGAGCCCGAGAGAGTGCCTCCTGAAGAGAGA | 8962 |
| QY | 182 | GCCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAA | 241 |
| Db | 8963 | GCCCAGAATGTCAACGGCTCCTACGAAAGCAAGCTGAGGAGGTCAACACTCAGTGGGAA | 9022 |
| QY | 242 | AAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTC | 301 |
| Db | 9023 | AAACTGAACCTGCACTCTGCAGACTGGCAGAGAAAAATAGACGAGGCCCTCGAAAGACTC | 9082 |
| QY | 302 | CAGGAACCTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATC | 361 |
| Db | 9083 | CAGGAGCTTCAGGAAGCAACAGATGAGCTGGATCTCAAACCTACGTAGGCAGAGGTGATC | 9142 |
| QY | 362 | AAGGGATCCTGGCAGCCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAG | 421 |
| Db | 9143 | AAAGGATCCTGGCAGCCCTGTGGGTGACCTCCTCATTTGACTCTCTCCAAGATCACCTCGAA | 9202 |
| QY | 422 | AAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTTGAAAGAGAACGTGAGCCACGTCAAT | 481 |
| Db | 9203 | AAAGTCAAGGCGCTTCGAGGAGAAATTACACCTCTGAAAGAGAAATGTACGTACGTCAAT | 9262 |
| QY | 482 | GACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCACT | 541 |
| Db | 9263 | GACCTTGCTCGCCAACTCACTACGTGGGCATTTCAGCTGTACCCATATAACCTCAACACT | 9322 |
| QY | 542 | CTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCGTTCGAGGACCGAGTCAGG | 601 |
| Db | 9323 | CTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCATTGAGGACCGCATCAGG | 9382 |
| QY | 602 | CAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTTCTTCCACGTC | 661 |
| Db | 9383 | CAGCTGCATGAAGCGCACAGGGACTTTGGACCAGCCTCCAGCACTTCTTTCACACTTCT | 9442 |
| QY | 662 | GTCCAGGCTCCCTGGGAGAGGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCAC | 721 |
| Db | 9443 | GTCCAGGCTCCCTGGGAGAGGCCATCTCACCACAAAGTGCCTTACTATATCAACCAC | 9502 |
| QY | 722 | GAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTTACCAGTCTTTAGCT | 781 |
| Db | 9503 | GAGACCCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTTACCAGTCTTTAGCT | 9562 |
| QY | 782 | GACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAG | 841 |
| Db | 9563 | GACCTGAATAATGTCAGGTTCTCAGCTTACAGGACTGCCATGAAACTCCGAAGACTGCAG | 9622 |
| QY | 842 | AAGGCCCTTTGCTTGGATCTCTTCAGCCTGTTCAGCTGCATGTGATGCCTTGGACCAGCAC | 901 |
| Db | 9623 | AAGGCCCTTTGCTTGGATCTCTTCAGCCTTACGGCTGCATGCGATGCCTTGGACCAGCAC | 9682 |
| QY | 902 | AACCTCAAGCAAAATGACCGCCCATGGATATCTTCAGATTATTAATTGTTTGACCACT | 961 |
| Db | 9683 | AACCTCAAGCAAAATGACCGCCCATGGATATCTTCAGGTCATTAATACTGTCTGACCAC | 9742 |
| QY | 962 | ATTTATGACCGCCTGGAGCAAGACACAACAAATTTGGTCAACGTCCCTCTCTGCGTGGAT | 1021 |
| Db | 9743 | ATTTATGATCGCCTAGAGCAAGACACAACAAATCTGGTCAACGTCCCTCTCTGCGTGGAT | 9802 |
| QY | 1022 | ATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGTC | 1081 |
| Db | 9803 | ATGTGCTCAATTGGCTGCTGAATGTTTATGACACGGGACGAAACGGGGAGGATCCGGGTC | 9862 |
| QY | 1082 | CTGTCTTTAAAAACTGGCATCAATTCCTGTGTAAAGCAATTTGGAAGACAAGTACAGA | 1141 |

Db 9863 CTGCTCTTTAAACTGGCATCAATTTCTGTGTAAAGCCCATTTGGAAGACAAGTACAGA 9922

QY 1142 TACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCTC 1201

Db 9923 TACCTCTTCAAGCAAGTGGCAAGTTGACAGGATTTTGTGACCAGCGCAGGCTGGGCCTC 9982

QY 1202 CTCTCTGCATGATTTCTATCCAAATTCACAAAGACAGTTGGGTGAAGTTGCATCCTTTTGGGGGC 1261

Db 9983 CTCTCTGCATGACTCTATCCAGATCCCAAGACAGTTGGGTGAAGTCCGATCGTTTCGGGGGC 10042

QY 1262 AGTAACATTTAGCCCAAGTGTCCGGAGCTGCTTCCAAATTTG 1301

Db 10043 AGTAACATTTAGCCCGAGTGTTCAGGAGCTGCTTCCAGTTTG 10082

RESULT 14

AX306153

LOCUS AX306153 13815 bp DNA linear PAT 11-DEC-2001

DEFINITION Sequence 904 from Patent WO0188188.

ACCESSION AX306153

VERSION AX306153.1 GI:17645441

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1

AUTHORS Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.

TITLE Method for examining ischemic conditions

JOURNAL Patent: WO 0188188-A 904 22-NOV-2001;

School Juridical Person Nihon University (JP)

FEATURES

source

1. .13815

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

BASE COUNT 4577 a 2719 c 3025 g 3494 t

ORIGIN

Query Match 87.3%; Score 1135.2; DB 6; Length 13815;

Best Local Similarity 92.1%; Pred. No. 0;

Matches 1197; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 2 GACTTTCCAGCAGTTTCAAGACGAGAACGATGACATAGGGCCTTCAAGAGGGAATTGAAA 61

Db 8708 GATTTCACGAGTTTCAAGACGAGAAATGATATACATAGGGCCTTCAAGAGGGAATTGAAA 8767

QY 62 ACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAG 121

Db 8768 ACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTGAGAATATTTCTGACAGAGCAG 8827

QY 122 CTTTGGAGGACTAGAGAAACTCTACGAGGAGCCCCAGAGAGCTGCCTCCTGAGGAGAGA 181

Db 8828 CTTTGGAGGACTAGAGAAACTCTACGAGGAGCCCCAGAGAACTGCCTCCTGAAGAAAGA 8887

QY 182 GCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAA 241

Db 8888 GCTCAGAAATGTCACTCGGCTCCTACGAAAGCAGGCTGAAGAGGTCAACGCTGAATGGGAC 8947

QY 242 AAATTGAACCTGCATCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAAGACTC 301

Db 8948 AAATTGAACCTGCGCTCAGCTGATTGGCAGAGAAAATAGATGAAGCTTTGAAAGACTC 9007

QY 302 CAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATC 361

Db 9008 CAGGAACCTTCAAGAGGCTGCGGATGAACTGGACCTCAAGTTGCGCCCAAGCTGAGGTGATC 9067

QY 362 AAGGATCTTGGCAGCCCCGTGGCGGATCTCTCAATGACTCTCTCCAGATCACTCGAG 421

Db 9068 AAGGATCTTGGCAGCCCCAGTGGGGGATCTCTCAATTGACTCTCTGCAAGATCACCTTGAA 9127

QY 422 AAAGTCAAGGCACCTTCGAGGAGAAAATTCGCGCTCTGAAAGAGAACGTGAGCCACGTCAAT 481

Db 9128 AAAGTCAAGGCACCTTCGGGGAGAAAATTGCACCTCTTAAAGAGAATGTCAATCGTGTCAAT 9187

QY 482 GACCTTGTCTGCCAGCTTACCACTTTTGGGCATTTACGCTCTCACCGTATATAACCTCAGCACT 541

Db 9188 GACCTTGCACATCAGCTGACCACTGGGCATTTACGCTCTCACCTTATAACCTCAGCACT 9247

QY 542 CTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAAG 601

Db 9248 TTGGAAGATCTGAATACCAAGATGAGGCTTCTACAGGTGGCTGTGGAGGACCGCTGTGAGA 9307

QY 602 CAGCTGCATGAAGCCCCACAGGGACTTTTGGTCCAGCATCTCAGCACATTTCTTTTCCAGTCT 661

Db 9308 CAGCTGCATGAAGCCCCACAGGGACTTTTGGTCCCTGCTGCAATCCAGCACATTTCCCTTCCACTTCA 9367

QY 662 GTCCAGGTCCTCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCAC 721

Db 9368 GTTCAGGTCCTCTGGGAGAGAGAGCCCATCTCACCAACAAAGTGCCTTACTATATCAACCAC 9427

QY 722 GAGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTTAGCT 781

Db 9428 GAGACCCAAACCACTTGTGTGGGACCAACCCCAAAATGACAGAGCTCTACCAGTCTTTTAGCT 9487

QY 782 GACCTGAATAATGTGAGTTCTCAGCTTATAGGACTGCGCATGAAACTCCGAAGACTGCGAG 841

Db 9488 GACCTGAATAATGTGAGTTCTCCTCCCGTATAGGACTGCCATGAAGCTCAGAAGGCTCCAG 9547

QY 842 AAGGCCCTTGTGTGGATCTCTTGAGCCTGTGAGCTGCTGATGCTGATGCTGATGCTGATGCTGAT 901

Db 9548 AAGGCCCTTGTGTGGATCTCTTGAGCCTGTGAGCTGCTGATGCTGATGCTGATGCTGATGCTGAT 9607

QY 902 AACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAATATTAATTGTTTGAACCACT 961

Db 9608 AACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAATATTAATTGTTTGAACCACT 9667

QY 962 ATTTATGACCGCTCTGGAGCAAGAGCACAAACAATTTGGTCAACGCTCCCTCTCTGCGTGGAT 1021

Db 9668 ATTTATGATCGTCTGGAGCAAGAGCACAAACAATCTGGTCAATGTCCCTCTCTGTTGGAT 9727

QY 1022 ATGTGTCTGAAGTGGCTGTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTC 1081

Db 9728 ATGTGTCTCAACTGGCTTCTCAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTC 9787

QY 1082 CTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGA 1141

Db 9788 CTGTCTTTTAAAACTGGCATCATTTCTGTGTAAAGCACACATTTGGAAGACAAGTACAGA 9847

QY 1142 TACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTACCAAGCGCAGGCTGGGCCTC 1201

Db 9848 TACCTTTTCAAGCAAGTGGCAAGTTTCAACTGGCTTTTGTGACCAGGTAGGCTGGGTCTT 9907

QY 1202 CTTCTGCATGATTTCTATCCAAATTCACAGACAGTTGGGTGAAGTTGCAATCCTTTGGGGGC 1261

Db 9908 CTTCTGCATGATTTCTATTCAAATCCCAAGACAGTTGGGTGAAGTTGCTTCCCTTTGGGGGC 9967

QY 1262 AGTAACATTTAGCCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301

Db 9968 AGTAACATTTAGCCCGAGTGTGAGGAGCTGCTTCCAATTTG 10007

RESULT 15

AX538582

LOCUS AX538582 13815 bp DNA linear PAT 23-NOV-2002

DEFINITION Sequence 2 from Patent WO0229056.

ACCESSION AX538582

VERSION AX538582.1 GI:25271088

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1

AUTHORS Chamberlain,J.S. and Harper,S.Q.

TITLE Mini-dystrophin nucleic acid and peptide sequences

JOURNAL Patent: WO 0229056-A 2 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES Location/Qualifiers
source 1.13815
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
BASE COUNT 4577 a 2719 c 3025 g 3494 t
ORIGIN

Query Match 87.3%; Score 1135.2; DB 6; Length 13815;
Best Local Similarity 92.1%; Pred. No. 0;
Matches 1197; Conservative 0; Mismatches 103; Indels 0; Gaps 0;
QY 2 GACTTTCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAGAGGGAAATTGAAA 61
Db 8708 GATTTCCAGCAGTTTCAGAAAGCAGAAATGATATACATAGGGCCTTCAGAGGGAAATTGAAA 8767
QY 62 ACTAAAGAACCTGTATCATGAGTACTCTTGGAGACTGTACGAATATTCTGACAGAGCAG 121
Db 8768 ACTAAAGAACCTGTATCATGAGTACTCTGGAGACTGTGAGAAATATTCTGACAGAGCAG 8827
QY 122 CCTTTGGAAGGACTAGAGAAACTCTACCAGAGGCCAGAGAGCTGCCCTCCTGAGGAGAGA 181
Db 8828 CCTTTGGAAGGACTAGAGAAACTCTACCAGAGGCCAGAGAACTGCCCTCCTGAAGAAAGA 8887
QY 182 GCCCAGAATGTCACCTGGCTTCTACGAAAGCAGGCTGAGGAGGTCACTAGTGGGAA 241
Db 8888 GCTCAGAATGTCACCTGGCTCCTACGAAAGCAGGCTGAAGAGGTCAACGCTGAATGGGAC 8947
QY 242 AAATTGAACCTGCATCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTC 301
Db 8948 AAATTGAACCTGCCTCAGCTGATTGGCAGAGAAAAATAGATGAAGCTCTTGAAAGACTC 9007
QY 302 CAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATC 361
Db 9008 CAGGAACCTTCAAGAGGTGCCGATGAACCTGGACCTCAAGTTGGCCCAAGCTGAGGTGATC 9067
QY 362 AAGGGATCCTGGCAGCCCGTGGCGGATCTCTCATATTGACTCTCTCCAAGATCACCTCGAG 421
Db 9068 AAGGGATCCTGGCAGCCAGTGGGGGATCTCCTCATTTGACTCTCTGCAAGATCACCTTGAA 9127
QY 422 AAAGTCAAGGCACCTTCGAGGAGAAAAATTGGCCCTCTGAAAGAGAACCTGAGCCACGTCAAT 481
Db 9128 AAAGTCAAGGCACCTTCGGGAGAAAAATTGCACCTCTTAAAGAGAAATGTCATCGTGTCAAT 9187
QY 482 GACCTTGCTCGCCAGCTTACCACCTTTTGGGCAATTGAGCTCTCACCGTATAACCTCAGCACT 541
Db 9188 GACCTTGACATCAGCTGACCCACACTGCGGCAATTGAGCTCTCACCTTATAACCTCAGCACT 9247
QY 542 CTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGG 601
Db 9248 TTGGAAGATCTGAATACAGATGGAGGCTTCTACAGGTGGCTGTGGAGGACCGTGTTCAGA 9307
QY 602 CAGCTGCATGAAGCCCCACAGGGACTTTTGGTCCAGCATCTCAGCACCTTCTTTCCACGTCT 661
Db 9308 CAGCTGCATGAAGCCCCACAGGGACTTTTGGTCCCTGCATCCAGCACCTTCTTCCACTTCA 9367
QY 662 GTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCCAAAACAAAGTGCCCTACTATATCAACCAC 721
Db 9368 GTTCAGGGTCCCTGGGAGAGAGCCCATCTCACCAAAACAAAGTGCCCTACTATATCAACCAC 9427
QY 722 CAGACTCAAACCAACTTGTGGGACCATATCCCAAAATGACAGAGCTCTACAGTCTTTAGCT 781
Db 9428 GAGACCCAAACCACTTTGTGGGACCAACCCCAAAATGACAGAGCTCTACAGTCTTTAGCT 9487
QY 782 GACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAG 841
Db 9488 GACCTGAATAATGTGAGATTCTCCGGGTATAGGACTGCCATGAAGCTCAGAAGGCTCCAG 9547
QY 842 AAGGCCCTTTGCTTGGATCTCTTGGACCTGTACGCTGCATGTGATGCTTGGACCAAGCAC 901
Db 9548 AAGGCCCTTTGCTTGGATCTCTTGGACCTGTACGCTGCATGTGATGCTTGGACCAAGCAC 9607

QY 902 AACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATATTAAATTGTTGACCACT 961
Db 9608 AACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATAAATTAAGTGTGACTACA 9667
QY 962 ATTTATGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGCTCCCTCTCTGCGTGGAT 1021
Db 9668 ATTTATGATCGTCTGGAGCAAGAGCACAACAATCTGGTCAATGTCCCTCTCTGTGTGGAT 9727
QY 1022 ATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTC 1081
Db 9728 ATGTGTCTCAACTGGCTTCTCAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTC 9787
QY 1082 CTGTCTTTTAAAACTGGCATCATTTCCCTCTGTAAAGCACATTTGGAAGACAAGTACAGA 1141
Db 9788 CTGTCTTTTAAAACTGGCATCATTTCTCTGTGTAAAGCACACCTTGGAAAGACAAGTACAGA 9847
QY 1142 TACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAGCGGAGGCTGGGCCTC 1201
Db 9848 TACCTTTTCAAGCAAGTGGCAAGTTTCAACTGGCTTTTGTGACCAGCGTAGGCTGGTCTT 9907
QY 1202 CTTCTGCATGATTCCTATCCAAATTCOAAGACAGTTGGGTGAAGTTGCATCCCTTTGGGGGC 1261
Db 9908 CTTCTGCATGATTCCTATTCAAATCCOAAGACAGTTGGGTGAAGTTGCTTCCCTTTGGGGGC 9967
QY 1262 AGTAACATTGAGCCCAAGTGTCCGGAGCTCGCTTCCAATTG 1301
Db 9968 AGTAACATTGAGCCGAGTGTGAGGAGCTGCTTCCAATTG 10007

Search completed: February 1, 2004, 23:51:08
Job time : 4637.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 10:13:25 ; Search time 309.023 Seconds
(without alignments)
11364.749 Million cell updates/sec

Title: US-09-845-416-10_COPY_1800_3100
Perfect score: 1301
Sequence: 1 cgactttccagcagttcaga.....ccggagctgcttccaatttg 1301

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 1301 | 100.0 | 1821 | 24 | Human dystrophin r |
| 2 | 1301 | 100.0 | 2169 | 24 | Human dystrophin r |
| 3 | 1301 | 100.0 | 3510 | 24 | Human dystrophin m |
| 4 | 1301 | 100.0 | 3531 | 24 | Human dystrophin m |
| 5 | 1301 | 100.0 | 3858 | 24 | Human dystrophin m |
| 6 | 1301 | 100.0 | 3999 | 24 | Human dystrophin m |
| 7 | 1301 | 100.0 | 4182 | 24 | Human dystrophin m |
| 8 | 1301 | 100.0 | 4476 | 24 | Adeno-associated v |

| | | | | | | |
|----|--------|-------|-------|----|----------|---------------------|
| 9 | 1301 | 100.0 | 4498 | 24 | AAD37258 | Adeno-associated v |
| 10 | 1301 | 100.0 | 4825 | 24 | AAD37257 | Adeno-associated v |
| 11 | 1301 | 100.0 | 4848 | 24 | AAD37263 | Adeno-associated v |
| 12 | 1301 | 100.0 | 4966 | 24 | AAD37256 | Adeno-associated v |
| 13 | 1301 | 100.0 | 4990 | 24 | AAD37262 | Adeno-associated v |
| 14 | 1301 | 100.0 | 5060 | 24 | AAD37264 | Adeno-associated v |
| 15 | 1301 | 100.0 | 5149 | 24 | AAD37255 | Adeno-associated v |
| 16 | 1301 | 100.0 | 5339 | 24 | ABK81998 | DNA encoding mini- |
| 17 | 1301 | 100.0 | 5462 | 24 | ABK81999 | DNA encoding mini- |
| 18 | 1301 | 100.0 | 5952 | 22 | AAD06794 | Human dystrophin g |
| 19 | 1301 | 100.0 | 8689 | 24 | ABK82000 | DNA encoding mini- |
| 20 | 1301 | 100.0 | 11058 | 24 | AAD37229 | Human dystrophin p |
| 21 | 1301 | 100.0 | 11241 | 24 | ABK82005 | CDNA encoding huma |
| 22 | 1301 | 100.0 | 11443 | 24 | ABK82002 | DNA encoding mini- |
| 23 | 1301 | 100.0 | 12923 | 10 | AAN90338 | Sequence of human |
| 24 | 1301 | 100.0 | 13957 | 24 | ABT10904 | Human breast cance |
| 25 | 1301 | 100.0 | 13957 | 24 | ABS69900 | Human dystrophin g |
| 26 | 1301 | 100.0 | 13957 | 24 | ABN95786 | Gene #2284 used to |
| 27 | 1301 | 100.0 | 13957 | 24 | ABK81959 | CDNA encoding huma |
| 28 | 1290 | 99.2 | 13977 | 24 | ABS70403 | Human bone remodel |
| 29 | 1135.2 | 87.3 | 13815 | 24 | ABK81960 | CDNA encoding mous |
| 30 | 1135.2 | 87.3 | 13815 | 24 | ABI99799 | Mouse ischaemic co |
| 31 | 1135.2 | 87.3 | 19307 | 17 | AAT27558 | Shuttle vector pAd |
| 32 | 1133.6 | 87.1 | 13815 | 19 | AAV18885 | Mus musculus dystro |
| 33 | 1085.2 | 83.4 | 4402 | 21 | AAZ48567 | A rod shortened dy |
| 34 | 1017.6 | 78.2 | 3446 | 24 | AAD37242 | Human dystrophin m |
| 35 | 1017.6 | 78.2 | 4414 | 24 | AAD37260 | Adeno-associated v |
| 36 | 1016.6 | 78.1 | 5417 | 24 | ABK81997 | DNA encoding mini- |
| 37 | 1016 | 78.1 | 1434 | 24 | AAD37243 | Human dystrophin r |
| 38 | 886 | 68.1 | 4402 | 21 | AAZ48568 | A rod shortened dy |
| 39 | 750.4 | 57.7 | 4075 | 21 | AAZ48569 | A rod shortened dy |
| 40 | 746.2 | 57.4 | 3747 | 21 | AAZ48566 | A rod shortened dy |
| 41 | 692.6 | 53.2 | 3163 | 21 | AAZ48571 | A rod shortened dy |
| 42 | 582.4 | 44.8 | 2005 | 25 | ABT33376 | NOVX DNA sequence |
| 43 | 580.8 | 44.6 | 2463 | 25 | ABT33375 | NOVX DNA sequence |
| 44 | 567.2 | 43.6 | 11096 | 24 | ABK81962 | CDNA encoding mous |
| 45 | 553.8 | 42.6 | 6045 | 18 | AAT74665 | Utrophin minigene. |

ALIGNMENTS

RESULT 1
AAD37241
ID AAD37241 standard; DNA; 1821 BP.
XX
AC AAD37241;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin rod, hinge and CR domain regions encoding DNA #2.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular

PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
XX
PS Example 1; Page 45-46; 7lpp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin DNA fragment encoding rods R22, R23 and R24, hinge H4 and
CC CR domain regions.
XX
SQ Sequence 2169 BP; 623 A; 529 C; 524 G; 493 T; 0 other;
Query Match 100.0%; Score 1301; DB 24; Length 2169;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGSCCTTCAAGAGGGAATTGAA 60
DB 451 CGACTTTCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGSCCTTCAAGAGGGAATTGAA 510
QY 61 AACTAAGAACCTGTAATCATGAGTACTCTTGAGACTGTPACGAATATTTCTGACAGAGCA 120
DB 511 AACTAAGAACCTGTAATCATGAGTACTCTTGAGACTGTPACGAATATTTCTGACAGAGCA 570
QY 121 GCCTTTGGAAGGACTAGAGAACTCTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG 180
DB 571 GCCTTTGGAAGGACTAGAGAACTCTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG 630
QY 181 AGCCAGAAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 240
DB 631 AGCCAGAAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 690
QY 241 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGACT 300
DB 691 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGACT 750
QY 301 CCAGGAACCTCAAGAGGCCACCGATGAGCTGGACCTCAAGTGCGCCAAGCTGAGGTGAT 360
DB 751 CCAGGAACCTCAAGAGGCCACCGATGAGCTGGACCTCAAGTGCGCCAAGCTGAGGTGAT 810
QY 361 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACTCGA 420
DB 811 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACTCGA 870
QY 421 GAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTTGAGCCACGTCAA 480
DB 871 GAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTTGAGCCACGTCAA 930
QY 481 TGACCTTGCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 540
DB 931 TGACCTTGCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 990
QY 541 TCTGGAAGACCTGAACACCAGATGGAGCTTCTGAGGTGSCGCTCGAGGACCGAGTCAG 600
DB 991 TCTGGAAGACCTGAACACCAGATGGAGCTTCTGAGGTGSCGCTCGAGGACCGAGTCAG 1050
QY 601 GCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
DB 1051 GCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 1110
QY 661 TGTCCAGGGTCCCTGGGAGAGAGGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA 720
DB 1111 TGTCCAGGGTCCCTGGGAGAGAGGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA 1170
QY 721 CGAGACTCAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780

DB 1171 CGAGACTCAAAACAACCTTGCTGGGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGC 1230
QY 781 TGACCTGAATAATGTGATTTCTCAGATTTCTAGGACTGCGCATGAAACTCCGAAGACTGCA 840
DB 1231 TGACCTGAATAATGTGATTTCTCAGATTTCTAGGACTGCGCATGAAACTCCGAAGACTGCA 1290
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCCTGTGATGCTTGGACCAAGCA 900
DB 1291 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCCTGTGATGCTTGGACCAAGCA 1350
QY 901 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCTCTGCAGATTATTAATTGTTGACCAC 960
DB 1351 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCTCTGCAGATTATTAATTGTTGACCAC 1410
QY 961 TATTTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGGGA 1020
DB 1411 TATTTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGGGA 1470
QY 1021 TATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 1080
DB 1471 TATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 1530
QY 1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
DB 1531 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1590
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGCCT 1200
DB 1591 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGCCT 1650
QY 1201 CCTTCTGCATGATTTCTATCCAAATTTCCAAAGACAGTTGGTGAAGTTGATCCTTTGGGGG 1260
DB 1651 CCTTCTGCATGATTTCTATCCAAATTTCCAAAGACAGTTGGTGAAGTTGATCCTTTGGGGG 1710
QY 1261 CAGTAACATTGAGCCAAGTGTCCGAGCTGCTTCCAAATTTG 1301
DB 1711 CAGTAACATTGAGCCAAGTGTCCGAGCTGCTTCCAAATTTG 1751
RESULT 3
AAD37240
ID AAD37240 standard; DNA; 3510 BP.
XX
AC AAD37240;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta3510.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a

PT dystrophin gene -
XX
PS Example 1; Page 51-52; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3510 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1 and R2), 8407-10227 (rods R23 and R24, hinge H4 and
CC CR domain) and 11047-11058 (dystrophin last 3 amino acids).
XX
SQ Sequence 3510 BP; 1073 A; 787 C; 828 G; 822 T; 0 other;
Query Match 100.0%; Score 1301; DB 24; Length 3510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTCCAGCAGTTCAGAAAGCAGAACGATGTACATAGGGCCCTCAAGAGGGAAATCAA 60
Db 1779 CGACTTCCAGCAGTTCAGAAAGCAGAACGATGTACATAGGGCCCTCAAGAGGGAAATCAA 1838
QY 61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAAATATTTCTGACAGAGCA 120
Db 1839 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAAATATTTCTGACAGAGCA 1898
QY 121 GCCTTTGGAAGGACTAGAGAAACTCTACAGAGAGCCAGAGAGTGCCTCCTGAGGAGAG 180
Db 1899 GCCTTTGGAAGGACTAGAGAAACTCTACAGAGAGCCAGAGAGTGCCTCCTGAGGAGAG 1958
QY 181 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGA 240
Db 1959 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGA 2018
QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 300
Db 2019 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 2078
QY 301 CCAGGAACCTTCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGAT 360
Db 2079 CCAGGAACCTTCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGAT 2138
QY 361 CAAGGGATCCTGGCAGCCCGTGGCGATCTCTCATTGACTCTCTCCAAGATCACCTCGA 420
Db 2139 CAAGGGATCCTGGCAGCCCGTGGCGATCTCTCATTGACTCTCTCCAAGATCACCTCGA 2198
QY 421 GAAAGTCAAGGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAA 480
Db 2199 GAAAGTCAAGGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAA 2258
QY 481 TGACCTTGCTCGCCAGCTTACCCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAC 540
Db 2259 TGACCTTGCTCGCCAGCTTACCCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAC 2318
QY 541 TCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 600
Db 2319 TCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 2378
QY 601 GCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
Db 2379 GCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 2438
QY 661 TGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 720
Db 2439 TGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 2498
QY 721 CGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTTAGC 780

Db 2499 CGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTAGC 2558
QY 781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
Db 2559 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 2618
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGCTGATGATGCCTTTGGACCAGCA 900
Db 2619 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGCTGATGATGCCTTTGGACCAGCA 2678
QY 901 CAACCTCAAGCAAAATGACCAGCCCATGGATATCTCTGACATTAATAATTGTTGACCCAC 960
Db 2679 CAACCTCAAGCAAAATGACCAGCCCATGGATATCTCTGACATTAATAATTGTTGACCCAC 2738
QY 961 TATTTATGACCGCCTGGAGCAAGACACAAATAATTTGGTCAACGTCCTCTCTGCGTGA 1020
Db 2739 TATTTATGACCGCCTGGAGCAAGACACAAATAATTTGGTCAACGTCCTCTCTGCGTGA 2798
QY 1021 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGT 1080
Db 2799 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGT 2858
QY 1081 CCTGTCTTTAAAAAAGTGGCATCAATTTCCCTGTGTAAAGCAATTTGGAAGACAAGTACAG 1140
Db 2859 CCTGTCTTTAAAAAAGTGGCATCAATTTCCCTGTGTAAAGCAATTTGGAAGACAAGTACAG 2918
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCT 1200
Db 2919 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCT 2978
QY 1201 CCTTCTGATGATTTATCCAAATTTCCAAAGACAGTTTGGTGAAGTTGCATCCTTTGGGG 1260
Db 2979 CCTTCTGATGATTTATCCAAATTTCCAAAGACAGTTTGGTGAAGTTGCATCCTTTGGGG 3038
QY 1261 CAGTAAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
Db 3039 CAGTAAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAATTTG 3079
RESULT 4
AAD37238
ID AAD37238 standard; DNA; 3531 BP.
XX
AC AAD37238;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta3531.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a

PT dystrophin gene -
XX PS Example 1; Page 50-51; 7lpp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3531 containing nucleotides 1-1341 (N-terminus,
CC hinge H1 and rod R1), 8059-10227 (rods R22, R23 and R24, hinge H4 and
CC CR domain) and 11047-11058 (dystrophin last 3 amino acids).
XX
SQ Sequence 3531 BP; 1071 A; 809 C; 824 G; 827 T; 0 other;
Query Match 100.0%; Score 1301; DB 24; Length 3531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCAGCAGTTCAGAGCAGAACGATGTACATAGGSCCTTCAAGAGGGAATTGAA 60
DB 1800 CGACTTTCAGCAGTTCAGAGCAGAACGATGTACATAGGSCCTTCAAGAGGGAATTGAA 1859
QY 61 AACTAAAGAACTGTAAATCATGACTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
DB 1860 AACTAAAGAACTGTAAATCATGACTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 1919
QY 121 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG 180
DB 1920 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG 1979
QY 181 AGCCCAGAAATGTCACTGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 240
DB 1980 AGCCCAGAAATGTCACTGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 2039
QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 300
DB 2040 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 2099
QY 301 CCAGGAACCTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT 360
DB 2100 CCAGGAACCTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT 2159
QY 361 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 420
DB 2160 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 2219
QY 421 GAAAGTCAAGGCACCTTCGAGGAGAGAAATTGGCCCTCTGAAAGAGAACGCTGAGCCACGTCAA 480
DB 2220 GAAAGTCAAGGCACCTTCGAGGAGAGAAATTGGCCCTCTGAAAGAGAACGCTGAGCCACGTCAA 2279
QY 481 TGACCTTGCTCGCCAGCTTACCCACTTTGGGCATTTCAGCTCTCACCGTATATACCTCAGCAC 540
DB 2280 TGACCTTGCTCGCCAGCTTACCCACTTTGGGCATTTCAGCTCTCACCGTATATACCTCAGCAC 2339
QY 541 TCTGGAAGACCTTGAACACCAGATGGAGCTTCTGACGGTGGCCGTCGAGGACCGAGTCAG 600
DB 2340 TCTGGAAGACCTTGAACACCAGATGGAGCTTCTGACGGTGGCCGTCGAGGACCGAGTCAG 2399
QY 601 GCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTC 660
DB 2400 GCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTC 2459
QY 661 TGTCCAGGGTCCCTGGGAGAGAGGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA 720
DB 2460 TGTCCAGGGTCCCTGGGAGAGAGGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA 2519
QY 721 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780

DB 2520 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 2579
QY 781 TGACCTGAATAATGTCTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
DB 2580 TGACCTGAATAATGTCTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 2639
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCTTGGACCAGCA 900
DB 2640 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCTTGGACCAGCA 2699
QY 901 CAACCTCAAGCAAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCAC 960
DB 2700 CAACCTCAAGCAAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCAC 2759
QY 961 TATTTATGACCGCCTGGAGCAAGAGACAAACAATTGGTCAACGTCCTCTCTGCGTGGGA 1020
DB 2760 TATTTATGACCGCCTGGAGCAAGAGACAAACAATTGGTCAACGTCCTCTCTGCGTGGGA 2819
QY 1021 TATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1080
DB 2820 TATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 2879
QY 1081 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAG 1140
DB 2880 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAG 2939
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 1200
DB 2940 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 2999
QY 1201 CCTTCTGCATGATTCTATCCAAATTCAAAGACAGTTGGTGAAGTTGCATCCTTTGGGG 1260
DB 3000 CCTTCTGCATGATTCTATCCAAATTCAAAGACAGTTGGTGAAGTTGCATCCTTTGGGG 3059
QY 1261 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTG 1301
DB 3060 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTG 3100
RESULT 5
AAD37237
ID AAD37237 standard; DNA; 3858 BP.
XX AAD37237;
AC AAD37237;
XX
DT 21-AUG-2002 (first entry)
XX
Human dystrophin minigene delta3849.
DE
XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX Homo sapiens.
OS
XX WO200183695-A2.
PN
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
PI Xiao X;
XX
WPI; 2002-049342/06.
DR
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a

PT dystrophin gene -
XX
PS Example 1; Page 48-49; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3849 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1, R2), 8059-10227 (rods R22, R23 and R24, hinge H4
CC and CR domain) and 11047-11058 (dystrophin last 3 amino acids).
XX
SQ Sequence 3858 BP; 1189 A; 866 C; 905 G; 898 T; 0 other;
Query Match 100.0%; Score 1301; DB 24; Length 3858;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCAGCAGTTCAGAAAGCAGAACGATGTACATAGGGCTTCAAGAGGGAAATTGAA 60
Db 2127 CGACTTTCAGCAGTTCAGAAAGCAGAACGATGTACATAGGGCTTCAAGAGGGAAATTGAA 2186
QY 61 AACTAAAGAACTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db 2187 AACTAAAGAACTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 2246
QY 121 GCCTTTTGGAGGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCCTCCTGAGGAGAG 180
Db 2247 GCCTTTTGGAGGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCCTCCTGAGGAGAG 2306
QY 181 AGCCGAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 240
Db 2307 AGCCGAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 2366
QY 241 AAAATTGAACCTTGCACTCCGCTGACTGCGAGAGAAAAATAGATGAGACCCCTTGAAAGACT 300
Db 2367 AAAATTGAACCTTGCACTCCGCTGACTGCGAGAGAAAAATAGATGAGACCCCTTGAAAGACT 2426
QY 301 CCAGGAACCTTCAGAGGCCACCGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 360
Db 2427 CCAGGAACCTTCAGAGGCCACCGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 2486
QY 361 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 420
Db 2487 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 2546
QY 421 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGCTGAGCCACGTCAA 480
Db 2547 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGCTGAGCCACGTCAA 2606
QY 481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTACCGTATAACCTCAGCAC 540
Db 2607 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTACCGTATAACCTCAGCAC 2666
QY 541 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAG 600
Db 2667 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAG 2726
QY 601 GCAGCTGCATGAAGCCACACAGGGACTTTGTGCCAGCATCTCAGCACTTTCTTTCCACGTC 660
Db 2727 GCAGCTGCATGAAGCCACACAGGGACTTTGTGCCAGCATCTCAGCACTTTCTTTCCACGTC 2786
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCA 720
Db 2787 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCA 2846
QY 721 CGAGACTCAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780

Db 2847 CGAGACTCAAACAACCTTGCTGGGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGC 2906
QY 781 TGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCCATGAAACTCCGAAGACTGCA 840
Db 2907 TGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCCATGAAACTCCGAAGACTGCA 2966
QY 841 GAAGGCCCTTTGCTTGGATCTCTGAGCCTGTGAGCTGTCATGTGATGCCCTTGGACCAGCA 900
Db 2967 GAAGGCCCTTTGCTTGGATCTCTGAGCCTGTGAGCTGTCATGTGATGCCCTTGGACCAGCA 3026
QY 901 CAACCTCAAGCAAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC 960
Db 3027 CAACCTCAAGCAAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC 3086
QY 961 TATTATGACCGCTGAGCAAGAGACACAACAATTTGGTCAACGTCCTCTCTGCGTGGGA 1020
Db 3087 TATTATGACCGCTGAGCAAGAGACACAACAATTTGGTCAACGTCCTCTCTGCGTGGGA 3146
QY 1021 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1080
Db 3147 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 3206
QY 1081 CCTGTCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAG 1140
Db 3207 CCTGTCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAG 3266
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGGCT 1200
Db 3267 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGGCT 3326
QY 1201 CCTTCTGCATGATTCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1260
Db 3327 CCTTCTGCATGATTCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 3386
QY 1261 CAGTAACATTGAGCCCAAGTGTCCGAGCTGCTTCCAATTG 1301
Db 3387 CAGTAACATTGAGCCCAAGTGTCCGAGCTGCTTCCAATTG 3427

RESULT 6

AAD37234
ID AAD37234 standard; DNA; 3999 BP.
XX
AC AAD37234;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta3990.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a

PT dystrophin gene -
XX
PS Example 1; Page 46-47; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3990 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1 and R2), 8059-10227 (rods R22, R23 and R24, hinge
CC H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids).
XX
SQ Sequence 3999 BP; 1223 A; 907 C; 933 G; 936 T; 0 other;

Query Match 100.0%; Score 1301; DB 24; Length 3999;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTTCCAGCAGTTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
Db |||||
QY 61 AACTAAAGAACTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db |||||
QY 121 GCCTTTGGAGACTAGAGAACTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 180
Db |||||
QY 2388 GCCTTTGGAGACTAGAGAACTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 2447
QY 181 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 240
Db |||||
QY 2448 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 2507
QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAATAATAGATGAGACCCCTTGAAAGACT 300
Db |||||
QY 301 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 360
Db |||||
QY 2568 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 2627
QY 361 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 420
Db |||||
QY 2628 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 2687
QY 421 GAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACTGAGGCCACGTCGCA 480
Db |||||
QY 2688 GAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACTGAGGCCACGTCGCA 2747
QY 481 TGACCTTGCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 540
Db |||||
QY 2748 TGACCTTGCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 2807
QY 541 TCTGGAAGACCTGAAACACCAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAG 600
Db |||||
QY 2808 TCTGGAAGACCTGAAACACCAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAG 2867
QY 601 GCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
Db |||||
QY 2868 GCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 2927
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 720
Db |||||
QY 2928 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 2987
QY 721 CGAGACTCAAACTGCTGTTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780

Db 2998 CGAGACTCAAACTTGTGGGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGC 3047
QY 781 TGACCTGAATAATGTCTAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCA 840
Db 3048 TGACCTGAATAATGTCTAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCA 3107
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGATGATGCTTGGACCAGCA 900
Db 3108 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGATGATGCTTGGACCAGCA 3167
QY 901 CAACCTCAAGCAAAATGACAGCCCATGGATATCTGCAGATTATTAATTTTGACCAC 960
Db 3168 CAACCTCAAGCAAAATGACAGCCCATGGATATCTGCAGATTATTAATTTTGACCAC 3227
QY 961 TATTTATGACCGCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTGCGTGA 1020
Db 3228 TATTTATGACCGCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTGCGTGA 3287
QY 1021 TATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1080
Db 3288 TATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 3347
QY 1081 CCTGTCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
Db 3348 CCTGTCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 3407
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGSCCT 1200
Db 3408 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGSCCT 3467
QY 1201 CCTTCTGTCATGATTTCTATCCAAATTCGAAGACAGTTGGTGAAGTTGCATCCTTTGGGG 1260
Db 3468 CCTTCTGTCATGATTTCTATCCAAATTCGAAGACAGTTGGTGAAGTTGCATCCTTTGGGG 3527
QY 1261 CAGTAACATTGAGCCCAAGTGTCCGAGAGCTGCTTCCAATTTG 1301
Db 3528 CAGTAACATTGAGCCCAAGTGTCCGAGAGCTGCTTCCAATTTG 3568

RESULT 7

AAD37230
ID AAD37230 standard; DNA; 4182 BP.

XX

AC AAD37230;

XX

DT 21-AUG-2002 (first entry)

XX

DE Human dystrophin minigene delta4173.

XX

KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

KW Becker muscular dystrophy; ds.

XX

OS Homo sapiens.

XX

PN WO200183695-A2.

XX

PD 08-NOV-2001.

XX

PF 27-APR-2001; 2001WO-US13677.

XX

PR 28-APR-2000; 2000US-200777P.

XX

PA (XIAO/) XIAO X.

XX

PI Xiao X;

XX

WPI; 2002-049342/06.

DR

XX New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,

PT rod repeats, H1 and H4 domains and a cysteine rich domain of a

PT dystrophin gene -

PS Example 1; Page 43-44; 7lpp; English.

XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta4173 containing nucleotides 1-1992 (N-terminus,
CC hinge H1 and rods R1, R2 and R3), 8059-10227 (rods R22, R23 and R24,
CC hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids).

XX
SQ Sequence 4182 BP; 1309 A; 927 C; 970 G; 976 T; 0 other;

Query Match 100.0%; Score 1301; DB 24; Length 4182;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTTCCAGCAGTTTCAGAAAGCAGACGATGTACATAGGSCCTTCAAGAGGGAAATTGAA 60
Db 2451 CGACTTTCCAGCAGTTTCAGAAAGCAGACGATGTACATAGGSCCTTCAAGAGGGAAATTGAA 2510
QY 61 AACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db 2511 AACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 2570
QY 121 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCGAGAGCTGCCTCCTGAGGAGAG 180
Db 2571 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCGAGAGCTGCCTCCTGAGGAGAG 2630
QY 181 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCATATCTGAGTGGGA 240
Db 2631 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCATATCTGAGTGGGA 2690
QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 300
Db 2691 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 2750
QY 301 CCAGGAACCTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 360
Db 2751 CCAGGAACCTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 2810
QY 361 CAAGGGATCCTGCGAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGA 420
Db 2811 CAAGGGATCCTGCGAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGA 2870
QY 421 GAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCCTCTGTAAAGAGAACGTGAGCCACGTCAA 480
Db 2871 GAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCCTCTGTAAAGAGAACGTGAGCCACGTCAA 2930
QY 481 TGACCTTGCTCGCCAGCTTACCACCTTTGGSCATTCAGCTCTCACCGTATAACCTCAGCAC 540
Db 2931 TGACCTTGCTCGCCAGCTTACCACCTTTGGSCATTCAGCTCTCACCGTATAACCTCAGCAC 2990
QY 541 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 600
Db 2991 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 3050
QY 601 GCAGCTGCATGAAGCCACAGGGACTTTTGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
Db 3051 GCAGCTGCATGAAGCCACAGGGACTTTTGTTCAGCATCTCAGCACTTTCTTTCCACGTC 3110
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 720
Db 3111 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 3170
QY 721 CGAGACTCAACAACTTGCTGGGACCAATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780

Db 3171 CGAGACTCAACAACTTGCTGGGACCAATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 3230
QY 781 TGACCTGAATAATGTCAAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
Db 3231 TGACCTGAATAATGTCAAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3290
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGTCATGTGATGCCCTTGGACAGCA 900
Db 3291 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGTCATGTGATGCCCTTGGACAGCA 3350
QY 901 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCTTCGAGATTATTAATTGTTTGACCAC 960
Db 3351 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCTTCGAGATTATTAATTGTTTGACCAC 3410
QY 961 TATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCCTCTCTGCGTGA 1020
Db 3411 TATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCCTCTCTGCGTGA 3470
QY 1021 TATGTCTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1080
Db 3471 TATGTCTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 3530
QY 1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAG 1140
Db 3531 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAG 3590
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTACCAGCGCAGGCTGGCCCT 1200
Db 3591 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTACCAGCGCAGGCTGGCCCT 3650
QY 1201 CCTTCTGCATGATTTCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 1260
Db 3651 CCTTCTGCATGATTTCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 3710
QY 1261 CAGTAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAATTG 1301
Db 3711 CAGTAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAATTG 3751

RESULT 8

AAD37259

ID AAD37259 standard; DNA; 4476 BP.

XX AAD37259;

AC AAD37259;

XX 21-AUG-2002 (first entry)

DT Adeno-associated virus vector plasmid, AAV-MCK-3510.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

XX Becker muscular dystrophy; ds.

OS Chimeric - Homo sapiens.

OS Chimeric - Unidentified.

XX WO200183695-A2.

PN 08-NOV-2001.

XX 27-APR-2001; 2001WO-US13677.

PF 28-APR-2000; 2000US-200777P.

XX (XIAO/) XIAO X.

PA Xiao X;

XX WPI; 2002-049342/06.

XX New dystrophin minigene for treating Duchenne or Becker muscular

XX dystrophy comprises an N-terminal domain or modified N-terminal domain,

PT

PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 63-65; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence.
XX
SQ Sequence 4476 BP; 1252 A; 1096 C; 1127 G; 1001 T; 0 other;

Query Match 100.0%; Score 1301; DB 24; Length 4476;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|------|--|------|
| QY | 1 | CGACTTTCCAGAGTTTCAGAAGCAGACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA | 60 |
| DB | 2535 | CGACTTTCCAGAGTTTCAGAAGCAGACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA | 2594 |
| QY | 61 | AACATAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGACCA | 120 |
| DB | 2595 | AACATAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGACCA | 2654 |
| QY | 121 | GCCTTTGGAGGACTAGAGAAACTCTACAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG | 180 |
| DB | 2655 | GCCTTTGGAGGACTAGAGAAACTCTACAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG | 2714 |
| QY | 181 | AGCCAGAAATGTCACCTGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGA | 240 |
| DB | 2715 | AGCCAGAAATGTCACCTGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGA | 2774 |
| QY | 241 | AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT | 300 |
| DB | 2775 | AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT | 2834 |
| QY | 301 | CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT | 360 |
| DB | 2835 | CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT | 2894 |
| QY | 361 | CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTATTGACTCTCTCCTCAAGATCACCTCGA | 420 |
| DB | 2895 | CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTATTGACTCTCTCCTCAAGATCACCTCGA | 2954 |
| QY | 421 | GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTTGAGCCACGTC | 480 |
| DB | 2955 | GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTTGAGCCACGTC | 3014 |
| QY | 481 | TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGC | 540 |
| DB | 3015 | TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGC | 3074 |
| QY | 541 | TCTGGAAGACCTGAAACACACAGATGGAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTC | 600 |
| DB | 3075 | TCTGGAAGACCTGAAACACACAGATGGAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTC | 3134 |
| QY | 601 | GCAGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGACATTTCTTCCACGTC | 660 |
| DB | 3135 | GCAGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGACATTTCTTCCACGTC | 3194 |
| QY | 661 | TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTCCCTACTATATCAACCA | 720 |
| DB | 3195 | TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTCCCTACTATATCAACCA | 3254 |
| QY | 721 | CGAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTAGC | 780 |

| | | | |
|----|------|--|------|
| DB | 3255 | CGAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC | 3314 |
| QY | 781 | TGACCTGAATAATGTCTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA | 840 |
| DB | 3315 | TGACCTGAATAATGTCTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA | 3374 |
| QY | 841 | GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCTTGGACCAGCA | 900 |
| DB | 3375 | GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCTTGGACCAGCA | 3434 |
| QY | 901 | CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCAC | 960 |
| DB | 3435 | CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCAC | 3494 |
| QY | 961 | TATTTATGACCGCCTGGAGCAAGAGACAAATTTGGTCAACGTCCTCTCTGCGTGA | 1020 |
| DB | 3495 | TATTTATGACCGCCTGGAGCAAGAGACAAATTTGGTCAACGTCCTCTCTGCGTGA | 3554 |
| QY | 1021 | TATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT | 1080 |
| DB | 3555 | TATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT | 3614 |
| QY | 1081 | CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG | 1140 |
| DB | 3615 | CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG | 3674 |
| QY | 1141 | ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGCCT | 1200 |
| DB | 3675 | ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGCCT | 3734 |
| QY | 1201 | CCTTCTGCATGATTCTATCCAAATTCACAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG | 1260 |
| DB | 3735 | CCTTCTGCATGATTCTATCCAAATTCACAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG | 3794 |
| QY | 1261 | CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG | 1301 |
| DB | 3795 | CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG | 3835 |

RESULT 9
AAD37258
ID AAD37258 standard; DNA; 4498 BP.
XX
AC AAD37258;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-3531.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Unidentified.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,

PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -

PS Example 1: Page 62-63; 71pp; English.

The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is AAV vector plasmid construct containing human dystrophin minigenes, a muscle creatine kinase (MCK) promoter and a small polyA signal sequence.

Sequence 4498 BP: 1251 A; 1118 C; 1123 G; 1006 T; 0 other;
SO

```
Query Match      100.0%; Score 1301; DB 24; Length 4498;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0;
```

Qy 1 CGACTTTCCAGCAGTTCAGAAGCGAGAACGATGTACATAGGGCCTTCAAGAGCGGAATTGAA 60
|||
Db 2557 CGACTTTCCAGCAGTTCAGAAGCGAGAACGATGTACATAGGGCCTTCAAGAGCGGAATTGAA 2616

Qy 61 AACTAAAGAACCTGTPAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
|||
Dδ 2617 AACTAAAGAACCTGTPAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 2676

QY 121 GCCTTTGGAAGGACTAGAGAAACTCTACAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAG 180

D6 2677 GCCTTTGGAAGGACTAGAGAAACTCTACAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAG 2736

QY 181 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 240
|||
Db 2737 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 2796

| | | | |
|----|------|---|------|
| Qy | 241 | AAAAATTGAACCTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT | 300 |
| | | | |
| Db | 2797 | AAAAATTGAACCTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT | 2856 |

QY 301 CCAGGAAC TTC AAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGTGAT 360

Db 2857 CCAGGAAC TTC AAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGTGAT 2916

QY

361 CAAGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 420
|||||

dB

2917 CAAGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 2976
|||||

QY 421 GAAAGTCAAGGCACCTTCGAGGAGAAAATTGGCCCTCTGAAAGAGAAACGTGAGCCACGTCAA 480
db 2977 GAAAGTCAAGGCACCTTCGAGGAGAAAATTGGCCCTCTGAAAGAGAAACGTGAGCCACGTCAA 3036

Qy 481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCAC 540
db 3037 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCAC 3096

QY 541 TCTGGAAGACCTGAACACCGATGGAAGCTTCTGCAGGTGGCGTTCGAGGACCGAGTCAG 600

pb 3097 TCTGGAAGACCTGAACACCGATGGAAGCTTCTGCAGGTGGCGTTCGAGGACCGAGTCAG 3156

| | | | |
|----|------|---|------|
| Qy | 601 | GCAGCTGCATGAAGCCACAGGACTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTC | 660 |
| Dδ | 3157 | GCAGCTGCATGAAGCCACAGGACTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTC | 3216 |

| Qy | 661 | TGTCAGGGTCCCTGGGAGAGGCCATCTCGCCAAACAAGTGCCTACTATATCAACCA | 720 |
|----|------|--|------|
| | | | |
| db | 3217 | TGTCAGGGTCCCTGGGAGAGGCCATCTCGCCAAACAAGTGCCTACTATATCAACCA | 3276 |

721 CGAGACTCAAACAACACTTGCTGGGACCATCCCAAATGACAGAGCTCTACCACTCTTAGC 780

| | | | |
|-----------|---|--|------|
| Db | 3277 | CGAGACTCAAACAACTTGTCTGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC | 3333 |
| QY | 781 | TGACCTGAATAATGTCTAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA | 840 |
| Db | 3337 | TGACCTGAATAATGTCTAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA | 3396 |
| QY | 841 | GAAGSCCCTTTGCTTGGATCTCTTGGACCTGTGAGCTGTCATGTGATGCCTTGGACCAAGCA | 900 |
| Db | 3397 | GAAGSCCCTTTGCTTGGATCTCTTGGACCTGTGAGCTGTCATGTGATGCCTTGGACCAAGCA | 3456 |
| QY | 901 | CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTTGTGACCAAC | 960 |
| Db | 3457 | CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTTGTGACCAAC | 3516 |
| QY | 961 | TATTTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGGA | 1020 |
| Db | 3517 | TATTTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGGA | 3576 |
| QY | 1021 | TATGTCTGTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT | 1080 |
| Db | 3577 | TATGTCTGTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT | 3636 |
| QY | 1081 | CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAGACAAGTACAG | 1140 |
| Db | 3637 | CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAGACAAGTACAG | 3696 |
| QY | 1141 | ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGAGGCTGGGCCT | 1200 |
| Db | 3697 | ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGAGGCTGGGCCT | 3756 |
| QY | 1201 | CCTTCTGCATGATTTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG | 1260 |
| Db | 3757 | CCTTCTGCATGATTTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG | 3816 |
| QY | 1261 | CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG | 1301 |
| Db | 3817 | CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG | 3857 |
| RESULT 10 | | | |
| AAD37257 | | | |
| ID | AAD37257 standard; DNA; 4825 BP. | | |
| XX | AAD37257; | | |
| AC | 21-AUG-2002 (first entry) | | |
| DT | Adeno-associated virus vector plasmid, AAV-MCK-delta3849. | | |
| XX | Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin; | | |
| KW | adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD; | | |
| KW | Becker muscular dystrophy; ds. | | |
| XX | Chimeric - Homo sapiens. | | |
| OS | Chimeric - Unidentified. | | |
| XX | WO200183695-A2. | | |
| PN | 08-NOV-2001. | | |
| XX | 27-APR-2001; 2001WO-US13677. | | |
| PF | 28-APR-2000; 2000US-200777P. | | |
| XX | (XIAO/) XIAO X. | | |
| PA | Xiao X; | | |
| PI | WPI; 2002-049342/06. | | |
| XX | New dystrophin minigene for treating Duchenne or Becker muscular | | |
| PT | dystrophy comprises an N-terminal domain or modified N-terminal domain, | | |

PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
XX Example 1; Page 61-62; 71pp; English.
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence.
XX
SQ Sequence 4825 BP; 1369 A; 1175 C; 1204 G; 1077 T; 0 other;
Query Match 100.0%; Score 1301; DB 24; Length 4825;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCCAGCAGTTCAAGAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
Db |||||
61 AACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db |||||
2944 AACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 3003
QY 121 GCCTTTGGAAGGACTAGAGAACTCTACCAGGAGCCCGAGAGCTGCCTCTGAGGAGAG 180
Db |||||
3004 GCCTTTGGAAGGACTAGAGAACTCTACCAGGAGCCCGAGAGCTGCCTCTGAGGAGAG 3063
QY 181 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 240
Db |||||
3064 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 3123
QY 241 AAAATTGAACCTGCCTCGCTGACTGGCGAGAGAAAAATAGATGAGACCTTGAAAGACT 300
Db |||||
3124 AAAATTGAACCTGCCTCGCTGACTGGCGAGAGAAAAATAGATGAGACCTTGAAAGACT 3183
QY 301 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAGCTGAGGTGAT 360
Db |||||
3184 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAGCTGAGGTGAT 3243
QY 361 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 420
Db |||||
3244 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 3303
QY 421 GAAAGTCAAGGCACCTTCGAGGAGAAATTTGCGCTCTGAAAGAGAACGTTGAGCCACGTCAA 480
Db |||||
3304 GAAAGTCAAGGCACCTTCGAGGAGAAATTTGCGCTCTGAAAGAGAACGTTGAGCCACGTCAA 3363
QY 481 TGACCTTGCTCGCCAGCTTACCACTTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAC 540
Db |||||
3364 TGACCTTGCTCGCCAGCTTACCACTTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAC 3423
QY 541 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGAGGTTGGCCGCTCGAGGACCGAGTCAG 600
Db |||||
3424 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGAGGTTGGCCGCTCGAGGACCGAGTCAG 3483
QY 601 GCAGCTGCATGAAGCCCAAGGGGACTTTTGGTCCAGCATCTCAGCACATTTCTTTCCACGTC 660
Db |||||
3484 GCAGCTGCATGAAGCCCAAGGGGACTTTTGGTCCAGCATCTCAGCACATTTCTTTCCACGTC 3543
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGGCCAAACAAAGTGCCCTACTATATCAACCA 720
Db |||||
3544 TGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGGCCAAACAAAGTGCCCTACTATATCAACCA 3603
QY 721 CGAGACTCAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780

Db 3604 CGAGACTCAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 3663
QY 781 TGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
Db |||||
3664 TGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3723
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGACAGTGCATGTGATGCCCTTGACCCAGCA 900
Db |||||
3724 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGACAGTGCATGTGATGCCCTTGACCCAGCA 3783
QY 901 CAACCTCAAGCAAAATGACAGCCCATGGATATCTCTGCAGATTATTAATTGTTGACCAC 960
Db |||||
3784 CAACCTCAAGCAAAATGACAGCCCATGGATATCTCTGCAGATTATTAATTGTTGACCAC 3843
QY 961 TATTTATGACCGCTTGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTGCGTGA 1020
Db |||||
3844 TATTTATGACCGCTTGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTGCGTGA 3903
QY 1021 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1080
Db |||||
3904 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 3963
QY 1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAG 1140
Db |||||
3964 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAG 4023
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCGGACGGCTGGGCCT 1200
Db |||||
4024 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCGGACGGCTGGGCCT 4083
QY 1201 CCTTCTGCATGATTCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1260
Db |||||
4084 CCTTCTGCATGATTCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 4143
QY 1261 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCCAATTTG 1301
Db |||||
4144 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCCAATTTG 4184
RESULT 11
AAD37263
ID AAD37263 standard; DNA; 4848 BP.
XX
AC AAD37263;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3849.
XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX Chimeric - Homo sapiens.
OS Chimeric - Cytomegalovirus.
XX Chimeric - Unidentified.
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS
XX Example 1; Page 68-70; 71pp; English.
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC cytomegalovirus (CMV) promoter and a small polyA signal sequence.
XX
SQ Sequence 4848 BP; 1405 A; 1144 C; 1180 G; 1119 T; 0 other;
Query Match 100.0%; Score 1301; DB 24; Length 4848;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCCAGCAGTTTCCAGAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAAATTGAA 60
Db
QY 61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db
QY 121 GCCTTTTGAAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG 180
Db
QY 3027 GCCTTTTGAAGGACTAGAGAACTCTACAGGAGCCAGAGAGTGCCTCCTGAGGAGAG 3086
QY 181 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGCGA 240
Db
QY 3087 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGCGA 3146
QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAAGACT 300
Db
QY 3147 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAAGACT 3206
QY 301 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 360
Db
QY 3207 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 3266
QY 361 CAAGGGATCCTGGAGCCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 420
Db
QY 3267 CAAGGGATCCTGGAGCCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 3326
QY 421 GAAAGTCAAGGCACTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAA 480
Db
QY 3327 GAAAGTCAAGGCACTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAA 3386
QY 481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAC 540
Db
QY 3387 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAC 3446
QY 541 TCTGGAAGACCTGAAACACACAGATGGAAGCTTCTGAGGTGGCGGTGAGGACCGAGTCAG 600
Db
QY 3447 TCTGGAAGACCTGAAACACACAGATGGAAGCTTCTGAGGTGGCGGTGAGGACCGAGTCAG 3506
QY 601 GCAGCTGCATGAAGCCACACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
Db
QY 3507 GCAGCTGCATGAAGCCACACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 3566
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
Db
QY 3567 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 3626

QY 721 CGAGACTCAAAACAACCTTGCTGGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
Db
QY 3627 CGAGACTCAAAACAACCTTGCTGGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGC 3686
QY 781 TGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
Db
QY 3687 TGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3746
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCTCAGCTGATGATGCCCTTGGACCAGCA 900
Db
QY 3747 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCTCAGCTGATGATGCCCTTGGACCAGCA 3806
QY 901 CAACCTCAAGCAAAANTGACAGCCCATGGATATCCTGCAAGATTATTAATTGTTTGACCAC 960
Db
QY 3807 CAACCTCAAGCAAAANTGACAGCCCATGGATATCCTGCAAGATTATTAATTGTTTGACCAC 3866
QY 961 TATTTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCCTCTCTGCGTGGGA 1020
Db
QY 3867 TATTTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCCTCTCTGCGTGGGA 3926
QY 1021 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 1080
Db
QY 3927 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 3986
QY 1081 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
Db
QY 3987 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 4046
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 1200
Db
QY 4047 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 4106
QY 1201 CCTTCTGTCATGATTTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1260
Db
QY 4107 CCTTCTGTCATGATTTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 4166
QY 1261 CAGTAACATTGAGCCCAAGTGTCCGAGCTGCTTCCAAATTTG 1301
Db
QY 4167 CAGTAACATTGAGCCCAAGTGTCCGAGCTGCTTCCAAATTTG 4207
RESULT 12
AAD37256
ID AAD37256 standard; DNA; 4966 BP.
XX
AC AAD37256;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-delta3990.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
OS Chimeric - Homo sapiens.
OS Chimeric - Unidentified.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -

XX
PS Example 1; Page 59-60; 71pp; English.

XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence.

XX
SQ Sequence 4966 BP; 1403 A; 1216 C; 1232 G; 1115 T; 0 other;

Query Match 100.0%; Score 1301; DB 24; Length 4966;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTTCCAGCACTTCAGAACGACGACATGTACATAGGGCCCTCAAGAGGGAATTGAA 60
Db |||||

QY 61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db |||||

QY 121 GCCTTTGGAAGGACTAGAGAACTCTACAGAGAGCCGAGAGCTGCCTCCTGAGGAGAG 180
Db |||||

QY 181 AGCCCAGAAATGTCATCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 240
Db |||||

QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAATAATAGATGAGACCCCTTGAAGACT 300
Db |||||

QY 301 CCAGGAACCTTCAAGAGGCCACGGATGAGCTTGGACCTCAAGCTGCGCAAGCTGAGGTGAT 360
Db |||||

QY 361 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 420
Db |||||

QY 421 GAAAGTCAAGGCCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTTGAGCCACGTCAA 480
Db |||||

QY 481 TGACCTTGCTCGGCAGCTTACCCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 540
Db |||||

QY 541 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTCAGGTGGCCGCTGAGGACCGAGTCAG 600
Db |||||

QY 601 GCAGCTGCATGAAGCCACAGGGACTTTTGTCTCAGCATCTCAGCACTTCTTTCCACGTC 660
Db |||||

QY 661 TGTCCAGGGTCCCTGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
Db |||||

QY 3685 TGTCCAGGGTCCCTGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 3744
Db |||||

QY 721 CGAGACTCAACAACACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
Db |||||

QY 781 TGACCTGAATAATGTCTAGATTCTCAGTTATAGGACTGCCATGAACTCCGAAGACTGCA 840
Db |||||

QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCTGATGTCCTTTGGACCAGCA 900
Db |||||

QY 901 CAACCTCAAGCAAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCAC 960
Db |||||

QY 961 TATTTATGACCGCCTGGAGCAAGAGACAAATTTGGTCAACGTCCTCTCTGCGTGGGA 1020
Db |||||

QY 1021 TATGTGCTGTAACCTGGCTGCTGAATGTTTATGATACGGGACGAAGGAGGATCCGTGT 1080
Db |||||

QY 1081 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAG 1140
Db |||||

QY 1141 ATACCTTTTTCAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGGAGGCTGGGCCT 1200
Db |||||

QY 1201 CCTTCTGCATGATCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 1260
Db |||||

QY 1261 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTTG 1301

QY 4285 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTTG 4325

RESULT 13

AAD37262

ID AAD37262 standard; DNA; 4990 BP.

AC AAD37262;

XX 21-AUG-2002 (first entry)

DT 21-AUG-2002 (first entry)

XX Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3990.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

XX Becker muscular dystrophy; ds.

OS Chimeric - Homo sapiens.

OS Chimeric - Cytomegalovirus.

OS Chimeric - Unidentified.

XX WO200183695-A2.

XX 08-NOV-2001.

XX 27-APR-2001; 2001WO-US13677.

XX 28-APR-2000; 2000US-200777P.

XX (XIAO/) XIAO X.

XX Xiao X;

XX WPI; 2002-049342/06.

XX

XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 70-71; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC muscle creatine kinase (MCK) enhancer, a cytomegalovirus (CMV) promoter
CC and a small polyA signal sequence.
XX
SQ Sequence 5060 BP; 1449 A; 1217 C; 1234 G; 1160 T; 0 other;
Query Match 100.0%; Score 1301; DB 24; Length 5060;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
2Y 1 CGACTTTCCAGCAGTTTCAGAAGCAGAACGATGTACATAGGCCCTTCAAGAGGGAATTGAA 60
Db 3119 CGACTTTCCAGCAGTTTCAGAAGCAGAACGATGTACATAGGCCCTTCAAGAGGGAATTGAA 3178
2Y 61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db 3179 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 3238
2Y 121 GCCTTTGGAAGGACTAGAGAACTCTACCAGGAGCCACAGAGAGCTGCCTCCTGAGGAGAG 180
Db 3239 GCCTTTGGAAGGACTAGAGAACTCTACCAGGAGCCACAGAGAGCTGCCTCCTGAGGAGAG 3298
2Y 181 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACGTAGTGGGA 240
Db 3299 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACGTAGTGGGA 3358
2Y 241 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAATAAGATGAGACCCTTGAAAGACT 300
Db 3359 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAATAAGATGAGACCCTTGAAAGACT 3418
2Y 301 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGAT 360
Db 3419 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGAT 3478
2Y 361 CAAGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 420
Db 3479 CAAGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 3538
2Y 421 GAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTTGAGCCACGTCAA 480
Db 3539 GAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTTGAGCCACGTCAA 3598
2Y 481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCCAGTCTCACCGTATACCTCAGCAC 540
Db 3599 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCCAGTCTCACCGTATACCTCAGCAC 3658
2Y 541 TCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 600
Db 3659 TCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 3718
2Y 601 GCAGCTGCATGAAGCCCCACAGGACCTTTGGTCCAGCATCTCAGCACCTTTCTTCCACGTC 660
Db 3719 GCAGCTGCATGAAGCCCCACAGGACCTTTGGTCCAGCATCTCAGCACCTTTCTTCCACGTC 3778
2Y 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720

Db 3779 TGTCCAGGTCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 3838
Qy 721 CGAGACTCAAAACAACTTGCTGGGACCAATCCCAAATGACAGAGCTCTACAGTCTTTAGC 780
Db 3839 CGAGACTCAAAACAACTTGCTGGGACCAATCCCAAATGACAGAGCTCTACAGTCTTTAGC 3898
Qy 781 TGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
Db 3899 TGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3958
Qy 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGTCATGTGATGCTTGGACCAGCA 900
Db 3959 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGTCATGTGATGCTTGGACCAGCA 4018
Qy 901 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTCGACATTTAATTGTTTGACCAC 960
Db 4019 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTCGACATTTAATTGTTTGACCAC 4078
Qy 961 TATTTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGGGA 1020
Db 4079 TATTTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGGGA 4138
Qy 1021 TATGTGTCTGAACTGGCTGTCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 1080
Db 4139 TATGTGTCTGAACTGGCTGTCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 4198
Qy 1081 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
Db 4199 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 4258
Qy 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 1200
Db 4259 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 4318
Qy 1201 CCTTCTGCATGATTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 1260
Db 4319 CCTTCTGCATGATTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 4378
Qy 1261 CAGTAACATTTGAGCCAAAGTGTCGGGAGCTGCTTCCAAATTG 1301
Db 4379 CAGTAACATTTGAGCCAAAGTGTCGGGAGCTGCTTCCAAATTG 4419
RESULT 15
AAD37255
ID AAD37255 standard; DNA; 5149 BP.
XX
AC AAD37255;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-delta4173.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Unidentified.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX

DR WPI; 2002-049342/06.
XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 57-59; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence.
XX
SQ Sequence 5149 BP; 1489 A; 1236 C; 1269 G; 1155 T; 0 other;
Query Match 100.0%; Score 1301; DB 24; Length 5149;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCCAGCAGTTCCAGAGCAGAACCGATGTACATAGGGCCTTCAAGAGGGGAATTGAA 60
Db
QY 61 AACTAAAGAACTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGCACAGACGA 120
Db
QY 121 GCCTTTGGAGGACTAGAGAACTCTACAGAGAGCCAGAGAGTGCCTCCTGAGGAGAG 180
Db
QY 3328 GCCTTTGGAGGACTAGAGAACTCTACAGAGAGCCAGAGAGTGCCTCCTGAGGAGAG 3387
QY 181 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATCTGAGTGGGA 240
Db
QY 3388 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATCTGAGTGGGA 3447
QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 300
Db
QY 3448 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 3507
QY 301 CCAGGAACCTCAAGAGGCCACCGATGAGTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 360
Db
QY 3508 CCAGGAACCTCAAGAGGCCACCGATGAGTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 3567
QY 361 CAAGGGATCCTGGGAGCCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 420
Db
QY 3568 CAAGGGATCCTGGGAGCCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 3627
QY 421 GAAAGTCAAGGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGCTGAGCCACGTCAA 480
Db
QY 3628 GAAAGTCAAGGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGCTGAGCCACGTCAA 3687
QY 481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAC 540
Db
QY 3688 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAC 3747
QY 541 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGCTGAGGACCGAGTCAAG 600
Db
QY 3748 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGCTGAGGACCGAGTCAAG 3807
QY 601 GCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTC 660
Db
QY 3808 GCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTC 3867
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720

Db 3868 TGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 3927
QY 721 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
Db 3928 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 3987
QY 781 TGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAAGACTGCA 840
Db 3988 TGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAAGACTGCA 4047
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGCAGCTGCATGTGATGCCCTTGGACCAACA 900
Db 4048 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGCAGCTGCATGTGATGCCCTTGGACCAACA 4107
QY 901 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAAC 960
Db 4108 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAAC 4167
QY 961 TATTATGACCGCCTGGAGCAAGAGCACAACAAATTTGGTCAACGTCCTCTCTGCGTGA 1020
Db 4168 TATTATGACCGCCTGGAGCAAGAGCACAACAAATTTGGTCAACGTCCTCTCTGCGTGA 4227
QY 1021 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1080
Db 4228 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 4287
QY 1081 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAG 1140
Db 4288 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAG 4347
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGGCTGGGCT 1200
Db 4348 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGGCTGGGCT 4407
QY 1201 CCTTCTGCATGATTCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 1260
Db 4408 CCTTCTGCATGATTCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 4467
QY 1261 CAGTAACATTGAGCCCAAGTGTCCGAGCTGCTTCCAATTG 1301
Db 4468 CAGTAACATTGAGCCCAAGTGTCCGAGCTGCTTCCAATTG 4508

Search completed: February 1, 2004, 12:04:33
Job time : 312.023 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 23:51:42 ; Search time 425.945 Seconds
(without alignments)
11131.886 Million cell updates/sec

Title: US-09-845-416-10_COPY_1800_3100
Perfect score: 1301
Sequence: 1 cgactttccagcagttcaga.....ccggagctggttccaatttg 1301

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2434939 seqs, 1822278265 residues

Total number of hits satisfying chosen parameters: 4869878

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB ID | Description |
|------------|-------|-------|--------|-------|------------------|
| 1 | 1301 | 100.0 | 1821 | 13 | US-09-845-416-13 |
| 2 | 1301 | 100.0 | 2169 | 13 | US-09-845-416-4 |
| 3 | 1301 | 100.0 | 3510 | 13 | US-09-845-416-12 |
| 4 | 1301 | 100.0 | 3531 | 13 | US-09-845-416-10 |
| 5 | 1301 | 100.0 | 3858 | 13 | US-09-845-416-9 |
| 6 | 1301 | 100.0 | 3999 | 13 | US-09-845-416-6 |
| 7 | 1301 | 100.0 | 4182 | 13 | US-09-845-416-2 |
| 8 | 1301 | 100.0 | 4476 | 13 | US-09-845-416-31 |
| 9 | 1301 | 100.0 | 4498 | 13 | US-09-845-416-30 |
| 10 | 1301 | 100.0 | 4825 | 13 | US-09-845-416-29 |
| 11 | 1301 | 100.0 | 4848 | 13 | US-09-845-416-35 |
| 12 | 1301 | 100.0 | 4966 | 13 | US-09-845-416-28 |
| 13 | 1301 | 100.0 | 4990 | 13 | US-09-845-416-34 |
| 14 | 1301 | 100.0 | 5060 | 13 | US-09-845-416-36 |
| 15 | 1301 | 100.0 | 5149 | 13 | US-09-845-416-27 |

| | | | | | | |
|------|--------|-------|-------|----|---------------------|-------------------|
| 16 | 1301 | 100.0 | 5339 | 13 | US-10-149-736-40 | Sequence 40, Appl |
| 17 | 1301 | 100.0 | 5462 | 13 | US-10-149-736-41 | Sequence 41, Appl |
| 18 | 1301 | 100.0 | 8689 | 13 | US-10-149-736-42 | Sequence 42, Appl |
| 19 | 1301 | 100.0 | 11058 | 13 | US-09-845-416-1 | Sequence 1, Appl |
| 20 | 1301 | 100.0 | 11443 | 13 | US-10-149-736-44 | Sequence 44, Appl |
| 21 | 1301 | 100.0 | 12057 | 13 | US-10-149-736-47 | Sequence 47, Appl |
| 22 | 1301 | 100.0 | 13957 | 10 | US-09-782-378A-22 | Sequence 22, Appl |
| 23 | 1301 | 100.0 | 13957 | 10 | US-09-880-107-2284 | Sequence 2284, Ap |
| 24 | 1301 | 100.0 | 13957 | 13 | US-10-149-736-1 | Sequence 1, Appl |
| 25 | 1301 | 100.0 | 14082 | 13 | US-10-341-434-108 | Sequence 108, App |
| 26 | 1135.2 | 87.3 | 13815 | 13 | US-10-149-736-2 | Sequence 2, Appl |
| 27 | 1017.6 | 78.2 | 3446 | 13 | US-09-845-416-14 | Sequence 14, Appl |
| 28 | 1017.6 | 78.2 | 4414 | 13 | US-09-845-416-32 | Sequence 32, Appl |
| 29 | 1016.6 | 78.1 | 5417 | 13 | US-10-149-736-39 | Sequence 39, Appl |
| 30 | 1016 | 78.1 | 1434 | 13 | US-09-845-416-15 | Sequence 15, Appl |
| 31 | 567.2 | 43.6 | 11096 | 13 | US-10-149-736-4 | Sequence 4, Appl |
| 32 | 553.8 | 42.6 | 10302 | 10 | US-09-782-378A-23 | Sequence 23, Appl |
| 33 | 553.8 | 42.6 | 10302 | 13 | US-10-149-736-3 | Sequence 3, Appl |
| 34 | 545 | 41.9 | 16531 | 13 | US-10-149-736-3 | Sequence 667, App |
| 35 | 472 | 36.3 | 887 | 13 | US-10-149-736-35 | Sequence 35, Appl |
| 36 | 324 | 24.9 | 324 | 13 | US-10-149-736-33 | Sequence 33, Appl |
| 37 | 289 | 22.2 | 387 | 13 | US-10-149-736-32 | Sequence 32, Appl |
| 38 | 216 | 16.6 | 216 | 13 | US-10-149-736-34 | Sequence 34, Appl |
| 39 | 114 | 8.8 | 114 | 13 | US-10-149-736-45 | Sequence 45, Appl |
| c 40 | 102.8 | 7.9 | 275 | 13 | US-10-029-386-16813 | Sequence 16813, A |
| c 41 | 102.8 | 7.9 | 587 | 13 | US-10-029-386-3113 | Sequence 3113, Ap |
| 42 | 54.2 | 4.2 | 449 | 11 | US-09-918-995-24084 | Sequence 24084, A |
| 43 | 54.2 | 4.2 | 2200 | 12 | US-10-108-260A-194 | Sequence 194, App |
| 44 | 54.2 | 4.2 | 2247 | 10 | US-09-960-253-157 | Sequence 157, App |
| 45 | 42.8 | 3.3 | 1690 | 13 | US-10-117-722-69 | Sequence 69, Appl |

ALIGNMENTS

RESULT 1
US-09-845-416-13
; Sequence 13, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver.. 2.1
; SEQ ID NO 13
; LENGTH: 1821
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-13

| | | | | | |
|-----------------------|-----|--|--------------|------------|--------------|
| Query Match | | 100.0%; | Score 1301; | DB 13; | Length 1821; |
| Best Local Similarity | | 100.0%; | Pred. No. 0; | | |
| Matches 1301; | | Conservative | 0; | Mismatches | 0; |
| | | | | Indels | 0; |
| | | | | Gaps | 0; |
| QY | 1 | CGACTTTCCAGCAGTTCCAGAACGACGATGTACATAGGGCCTTCAAGAGGGAATTGAA | 60 | | |
| DB | 103 | CGACTTTCCAGCAGTTCCAGAACGACGATGTACATAGGGCCTTCAAGAGGGAATTGAA | 162 | | |
| QY | 61 | AACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGATATTTCTGACAGAGCA | 120 | | |
| DB | 163 | AACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGATATTTCTGACAGAGCA | 222 | | |
| QY | 121 | GCCTTTGGAAGGACTAGAGAACTCTACAGAGGCCAGAGAGCTGCCTCTGAGGAGAG | 180 | | |
| DB | 223 | GCCTTTGGAAGGACTAGAGAACTCTACAGAGGCCAGAGAGCTGCCTCTGAGGAGAG | 282 | | |
| QY | 181 | AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA | 240 | | |

Db 283 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCAATACTAGTGGGA 342
QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT 300
Db 343 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT 402
QY 301 CCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 360
Db 403 CCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 462
QY 361 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 420
Db 463 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 522
QY 421 GAAAGTCAAGGCACTTCGAGGAGAAATGGCCCTCTGAAAGAGAACTGAGCCACGTCAA 480
Db 523 GAAAGTCAAGGCACTTCGAGGAGAAATGGCCCTCTGAAAGAGAACTGAGCCACGTCAA 582
QY 481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 540
Db 583 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 642
QY 541 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 600
Db 643 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 702
QY 601 GCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
Db 703 GCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 762
QY 661 TGTCCAGGGTCCCTGGGAGAGAGGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACCA 720
Db 763 TGTCCAGGGTCCCTGGGAGAGAGGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACCA 822
QY 721 CGAGACTCAACAACTTGCTGGGACCATATCCCAAAATGACAGAGCTCTACCACTTTTAGC 780
Db 823 CGAGACTCAACAACTTGCTGGGACCATATCCCAAAATGACAGAGCTCTACCACTTTTAGC 882
QY 781 TGACCTGAATAATGTGATGATCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCA 840
Db 883 TGACCTGAATAATGTGATGATCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCA 942
QY 841 GAAGGCCCTTTGCTTGATCTCTTGAGCCTGTGAGCTGCAATGTGATGCCCTTGACCAAGCA 900
Db 943 GAAGGCCCTTTGCTTGATCTCTTGAGCCTGTGAGCTGCAATGTGATGCCCTTGACCAAGCA 1002
QY 901 CAACCTCAAGCAAAATGACACAGCCCATGGATATCCTGCAGATTAATTAATGTTGACCCAC 960
Db 1003 CAACCTCAAGCAAAATGACACAGCCCATGGATATCCTGCAGATTAATTAATGTTGACCCAC 1062
QY 961 TATTTATGACCGCCTGGAGCAAGAGCACAACTTTGGTCAACGTCCCTCTCTGCGTGGGA 1020
Db 1063 TATTTATGACCGCCTGGAGCAAGAGCACAACTTTGGTCAACGTCCCTCTCTGCGTGGGA 1122
QY 1021 TATGTGCTGAACCTGGCTGCTGAATGTTATGATACGGGACGAACAGGAGGATCCGTGT 1080
Db 1123 TATGTGCTGAACCTGGCTGCTGAATGTTATGATACGGGACGAACAGGAGGATCCGTGT 1182
QY 1081 CCTGTCTTTTAAACTGSGATCATTTCCCTGTGTAAGACACATTTGGAAGACAAAGTACAG 1140
Db 1183 CCTGTCTTTTAAACTGSGATCATTTCCCTGTGTAAGACACATTTGGAAGACAAAGTACAG 1242
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGAGCGGAGGCTGGGCT 1200
Db 1243 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGAGCGGAGGCTGGGCT 1302
QY 1201 CCTTCTGCATGATTTCTATCCAAATCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 1260
Db 1303 CCTTCTGCATGATTTCTATCCAAATCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 1362
QY 1261 CAGTAACTTGAGCCAAAGTTCGGAGAGCTGCTTCCAAATTTG 1301

Db 1363 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1403
RESULT 2
US-09-845-416-4
; Sequence 4, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2169
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-4
Query Match 100.0%; Score 1301; DB 13; Length 2169;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCCAGCAGTTTCAGAAAGCAGATGTACATAGGCGCTTCAAGAGGGAATTGAA 60
Db 451 CGACTTTCCAGCAGTTTCAGAAAGCAGATGTACATAGGCGCTTCAAGAGGGAATTGAA 510
QY 61 AACTAAAGAACTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db 511 AACTAAAGAACTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 570
QY 121 GCCTTTGGAAAGGACTAGAGAAACTCTACCAAGAGCCAGAGAGTGCCTCCTGAGGAGAG 180
Db 571 GCCTTTGGAAAGGACTAGAGAAACTCTACCAAGAGCCAGAGAGTGCCTCCTGAGGAGAG 630
QY 181 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGA 240
Db 631 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGA 690
QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT 300
Db 691 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT 750
QY 301 CCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 360
Db 751 CCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 810
QY 361 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 420
Db 811 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 870
QY 421 GAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACTGAGCCACGTCAA 480
Db 871 GAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACTGAGCCACGTCAA 930
QY 481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 540
Db 931 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 990
QY 541 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 600
Db 991 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 1050
QY 601 GCAGCTGCATGAAGCCCAAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
Db 1051 GCAGCTGCATGAAGCCCAAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 1110
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACCA 720

Db 1111 TGTCAGGTCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 1170
QY 721 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
Db 1171 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 1230
QY 781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
Db 1231 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 1290
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGTCATGTCCTTGACCCAGCA 900
Db 1291 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGTCATGTCCTTGACCCAGCA 1350
QY 901 CAACCTCAAGCAAAATGACCGCCATGGATATCCTGCAGATTATTAAATTGTTGACCCAC 960
Db 1351 CAACCTCAAGCAAAATGACCGCCATGGATATCCTGCAGATTATTAAATTGTTGACCCAC 1410
QY 961 TATTTATGACCGCCTGGAGCAAGACAAACAATTTGGTCAACGTCCTCTCTGCGTGA 1020
Db 1411 TATTTATGACCGCCTGGAGCAAGACAAACAATTTGGTCAACGTCCTCTCTGCGTGA 1470
QY 1021 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 1080
Db 1471 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 1530
QY 1081 CCTGTCTTTTAAAACTGGCATCAATTCCTGTGTAAAGCACATTTGGAAGACAACTACAG 1140
Db 1531 CCTGTCTTTTAAAACTGGCATCAATTCCTGTGTAAAGCACATTTGGAAGACAACTACAG 1590
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGCGCTGGGCCT 1200
Db 1591 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGCGCTGGGCCT 1650
QY 1201 CCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1260
Db 1651 CCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1710
QY 1261 CAGTAACATTGAGCCAAAGTGTCGGAGCTGCTTCCAAATTTG 1301
Db 1711 CAGTAACATTGAGCCAAAGTGTCGGAGCTGCTTCCAAATTTG 1751

RESULT 3
US-09-845-416-12
; Sequence 12, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 3510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-12

Query Match 100.0%; Score 1301; DB 13; Length 3510;
Best local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCCAGCAGTTTCAGAAGCAGACGATGTACATAGGGCTTCAAGAGGGAATTGAA 60
Db 1779 CGACTTTCCAGCAGTTTCAGAAGCAGACGATGTACATAGGGCTTCAAGAGGGAATTGAA 1838

QY 61 AACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db 1839 AACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 1898
QY 121 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCCGAGAGAGCTGCCTCCTGAGGAGAG 180
Db 1899 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCCGAGAGAGCTGCCTCCTGAGGAGAG 1958
QY 181 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTTCAATACTGAGTGGGA 240
Db 1959 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTTCAATACTGAGTGGGA 2018
QY 241 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATPAGATGAGACCCCTTGAAAGACT 300
Db 2019 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATPAGATGAGACCCCTTGAAAGACT 2078
QY 301 CCAGGAACCTTCAAGAGGCCACCGGATGAGCTGAGCCTCAAGCTGCGCCAAAGCTGAGGTGAT 360
Db 2079 CCAGGAACCTTCAAGAGGCCACCGGATGAGCTGAGCCTCAAGCTGCGCCAAAGCTGAGGTGAT 2138
QY 361 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 420
Db 2139 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 2198
QY 421 GAAAGTCAAGGCACCTTCGAGGAGAAAATTCGCGCTCTGAAAGAGAACTGAGCCACGTCAA 480
Db 2199 GAAAGTCAAGGCACCTTCGAGGAGAAAATTCGCGCTCTGAAAGAGAACTGAGCCACGTCAA 2258
QY 481 TGACCTTGCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAAACCCTCAGCAC 540
Db 2259 TGACCTTGCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAAACCCTCAGCAC 2318
QY 541 TCTGGAAGACCTGAAACACCCAGATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAAG 600
Db 2319 TCTGGAAGACCTGAAACACCCAGATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAAG 2378
QY 601 GCAGCTGCATGAAGCCCAAGGAGACTTTGGTCCAGCATCTCAGCACCTTCTTCTCCACGTC 660
Db 2379 GCAGCTGCATGAAGCCCAAGGAGACTTTGGTCCAGCATCTCAGCACCTTCTTCTCCACGTC 2438
QY 661 TGTCCAGGTCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 720
Db 2439 TGTCCAGGTCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 2498
QY 721 CGAGACTCAAAACAACCTTGCTGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
Db 2499 CGAGACTCAAAACAACCTTGCTGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 2558
QY 781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
Db 2559 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 2618
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCCCTTGACCCAGCA 900
Db 2619 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCCCTTGACCCAGCA 2678
QY 901 CAACCTCAAGCAAAATGACCCAGCCATGGATATCCTGCAGATTATTAAATTGTTTGACCCAC 960
Db 2679 CAACCTCAAGCAAAATGACCCAGCCATGGATATCCTGCAGATTATTAAATTGTTTGACCCAC 2738
QY 961 TATTTATGACCGCCTGGAGCAAGACCAACAATTTGGTCAACGTCCTCTCTGCGTGA 1020
Db 2739 TATTTATGACCGCCTGGAGCAAGACCAACAATTTGGTCAACGTCCTCTCTGCGTGA 2798
QY 1021 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCCGTGT 1080
Db 2799 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCCGTGT 2858
QY 1081 CCTGTCTTTTAAAACTGGCATCAATTCCTGTGTAAAGCACATTTGGAAGACAACTACAG 1140
Db 2859 CCTGTCTTTTAAAACTGGCATCAATTCCTGTGTAAAGCACATTTGGAAGACAACTACAG 2918
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGGCCT 1200

| | | | | | |
|-----------------------|------|---|---------------|-----------|--------------|
| Query Match | | 100.0%; | Score 1301; | DB 13; | Length 3858; |
| Best Local Similarity | | 100.0%; | Pred. No. 0; | | |
| Matches 1301; | | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| QY | 1 | CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA | 60 | | |
| Db | 2127 | CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA | 2186 | | |
| QY | 61 | AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA | 120 | | |
| Db | 2187 | AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA | 2246 | | |
| QY | 121 | GCCTTTTGGAAAGSACTAGAGAACTCTACAGGAGCCCCAGAGAGTGCCTCCTTGAGGAGAG | 180 | | |
| Db | 2247 | GCCTTTTGGAAAGSACTAGAGAACTCTACAGGAGCCCCAGAGAGTGCCTCCTTGAGGAGAG | 2306 | | |
| QY | 181 | AGCCCAGAAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA | 240 | | |
| Db | 2307 | AGCCCAGAAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA | 2366 | | |
| QY | 241 | AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT | 300 | | |
| Db | 2367 | AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT | 2426 | | |
| QY | 301 | CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGTGCGCCCAAGCTGAGGTGAT | 360 | | |
| Db | 2427 | CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGTGCGCCCAAGCTGAGGTGAT | 2486 | | |
| QY | 361 | CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA | 420 | | |
| Db | 2487 | CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA | 2546 | | |
| QY | 421 | GAAAGTCAAGGCACTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAA | 480 | | |
| Db | 2547 | GAAAGTCAAGGCACTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAA | 2606 | | |
| QY | 481 | TGACCTTGCTCGCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC | 540 | | |
| Db | 2607 | TGACCTTGCTCGCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC | 2666 | | |
| QY | 541 | TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAG | 600 | | |
| Db | 2667 | TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAG | 2726 | | |
| QY | 601 | GCAGCTGCATGAAGCCACAGGGACTTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC | 660 | | |
| Db | 2727 | GCAGCTGCATGAAGCCACAGGGACTTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC | 2786 | | |
| QY | 661 | TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA | 720 | | |
| Db | 2787 | TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA | 2846 | | |
| QY | 721 | CGAGACTCAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTTAGC | 780 | | |
| Db | 2847 | CGAGACTCAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTTAGC | 2906 | | |
| QY | 781 | TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA | 840 | | |
| Db | 2907 | TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA | 2966 | | |
| QY | 841 | GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCTTGGACCAAGCA | 900 | | |
| Db | 2967 | GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCTTGGACCAAGCA | 3026 | | |
| QY | 901 | CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC | 960 | | |
| Db | 3027 | CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC | 3086 | | |
| QY | 961 | TATTTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGGGA | 1020 | | |
| Db | 3087 | TATTTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGGGA | 3146 | | |

| | | | |
|--|------|--|------|
| QY | 1021 | TATGTGTCGAACTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCCTGT | 1080 |
| Db | 3147 | TATGTGTCGAACTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCCTGT | 3206 |
| QY | 1081 | CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG | 1140 |
| Db | 3207 | CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG | 3266 |
| QY | 1141 | ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCCT | 1200 |
| Db | 3267 | ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCCT | 3326 |
| QY | 1201 | CCTTCTGCATGATCTATCCAAATTCAGAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG | 1260 |
| Db | 3327 | CCTTCTGCATGATCTATCCAAATTCAGAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG | 3386 |
| QY | 1261 | CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTTG | 1301 |
| Db | 3387 | CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTTG | 3427 |
| RESULT 6 | | | |
| US-09-845-416-6 | | | |
| ; Sequence 6, Application US/09845416 | | | |
| ; Publication No. US20030171312A1 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: XIAO, XIAO | | | |
| ; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE | | | |
| ; FILE REFERENCE: DE1142 | | | |
| ; CURRENT APPLICATION NUMBER: US/09/845,416 | | | |
| ; CURRENT FILING DATE: 2001-04-30 | | | |
| ; PRIOR APPLICATION NUMBER: 60/200,777 | | | |
| ; PRIOR FILING DATE: 2000-04-28 | | | |
| ; NUMBER OF SEQ ID NOS: 36 | | | |
| ; SOFTWARE: PatentIn Ver. 2.1 | | | |
| ; SEQ ID NO 6 | | | |
| ; LENGTH: 3999 | | | |
| ; TYPE: DNA | | | |
| ; ORGANISM: Homo sapiens | | | |
| US-09-845-416-6 | | | |
| Query Match 100.0%; Score 1301; DB 13; Length 3999; | | | |
| Best Local Similarity 100.0%; Pred. No. 0; | | | |
| Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | |
| QY | 1 | CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA | 60 |
| Db | 2268 | CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA | 2327 |
| QY | 61 | AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA | 120 |
| Db | 2328 | AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA | 2387 |
| QY | 121 | GCCTTTTGGAAAGSACTAGAGAACTCTACAGAGAGCCCCAGAGAGTGCCTCCTGAGGAGAG | 180 |
| Db | 2388 | GCCTTTTGGAAAGSACTAGAGAACTCTACAGAGAGCCCCAGAGAGTGCCTCCTGAGGAGAG | 2447 |
| QY | 181 | AGCCCAGAAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA | 240 |
| Db | 2448 | AGCCCAGAAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA | 2507 |
| QY | 241 | AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT | 300 |
| Db | 2508 | AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT | 2567 |
| QY | 301 | CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT | 360 |
| Db | 2568 | CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT | 2627 |
| QY | 361 | CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA | 420 |
| Db | 2628 | CAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA | 2687 |

QY 421 GAAAGTCAAGGCACTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTCGAGCCACGTCAA 480
Db |||||
QY 481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAAACCTCAGCAC 540
Db |||||
QY 541 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCGTCGAGGACCGAGTCAG 600
Db |||||
QY 601 GCAGCTGCATGAAGCCCCACAGGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
Db |||||
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
Db |||||
QY 721 CGAGACTCAAAACAACTTGCTGGGACCAATCCCAAAATGACAGAGTCTACCACTCTTTAGC 780
Db |||||
QY 781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
Db |||||
QY 841 GAAGGCCCTTTGCTGGATCTCTTGAGCCCTGTGAGTGCATGTGATGCCCTTGACCCAGCA 900
Db |||||
QY 901 CAACCTCAAGCAAAATGACAGCCCATGGATATCTGCAGATTATTAATGTTTGACCAC 960
Db |||||
QY 961 TATTATGACCGCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTGCGTGA 1020
Db |||||
QY 1021 TATGTGCTGAACCTGCTGTAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1080
Db |||||
QY 1081 CCTGTCTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGAAGACAAGTACAG 1140
Db |||||
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGCAGGCTGGGCCT 1200
Db |||||
QY 1201 CCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 1260
Db |||||
QY 1261 CAGTAACATTGAGCCAAAGTCTCCGAGCTGCTTCCAATTG 1301
Db |||||
QY 1301 CAGTAACATTGAGCCAAAGTCTCCGAGCTGCTTCCAATTG 1358
Db |||||

RESULT 7
US-09-845-416-2
; Sequence 2, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DB1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30

; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4182
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-2

Query Match 100.0%; Score 1301; DB 13; Length 4182;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGSCCTTCAAGAGGGAATTGAA 60
Db |||||
QY 61 AACTAAAGAACTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db |||||
QY 121 GCCTTTTGGAAAGGACTAGAGAAACTCTACGAGGAGCCAGAGAGTGCCTCTCTGAGGAGAG 180
Db |||||
QY 181 AGCCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 240
Db |||||
QY 241 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACT 300
Db |||||
QY 301 CCAGGAACTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT 360
Db |||||
QY 361 CAAGGGATCCTGGCAGCCCCGTTGGCGGATCTCTCATTTGACTCTCTCCAAGATCACCTCGA 420
Db |||||
QY 421 GAAAGTCAAGGCACTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTCGAGCCACGTCAC 480
Db |||||
QY 481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCAC 540
Db |||||
QY 541 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 600
Db |||||
QY 601 GCAGCTGCATGAAGCCCCACAGGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
Db |||||
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
Db |||||
QY 721 CGAGACTCAAAACAACTTGTGGACCAATCCCAAAATGACAGAGTCTTACCACTCTTTAGC 780
Db |||||
QY 781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
Db |||||
QY 841 GAAGGCCCTTTGCTGGATCTCTTGAGCCCTGTGAGTGCATGTGATGCTTGACCCAGCA 900
Db |||||
QY 901 CAAGGCCCTTTGCTGGATCTCTTGAGCCCTGTGAGTGCATGTGATGCCCTTGACCCAGCA 960
Db |||||
QY 961 TATTATGACCGCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTGCGTGA 1020
Db |||||
QY 1021 TATGTGCTGAACCTGCTGTAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1080
Db |||||
QY 1081 CCTGTCTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGAAGACAAGTACAG 1140
Db |||||
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGCAGGCTGGGCCT 1200
Db |||||
QY 1201 CCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 1260
Db |||||
QY 1261 CAGTAACATTGAGCCAAAGTCTCCGAGCTGCTTCCAATTG 1301
Db |||||
QY 1301 CAGTAACATTGAGCCAAAGTCTCCGAGCTGCTTCCAATTG 1358
Db |||||

QY 901 CAACCTCAAGCAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC 960
| | | | |
Db 3351 CAACCTCAAGCAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC 3410
| | | | |
QY 961 TATTTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGGA 1020
| | | | |
Db 3411 TATTTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGGA 3470
| | | | |
QY 1021 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTTG 1080
| | | | |
Db 3471 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTTG 3530
| | | | |
QY 1081 CCTGTCTTTTAAACTGGCATCATTTTCCCTGTGTAAAGCACATTTTGAAGACAAAGTACAG 1140
| | | | |
Db 3531 CCTGTCTTTTAAACTGGCATCATTTTCCCTGTGTAAAGCACATTTTGAAGACAAAGTACAG 3590
| | | | |
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCGCT 1200
| | | | |
Db 3591 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCGCT 3650
| | | | |
QY 1201 CCTTCTGCATGATCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 1260
| | | | |
Db 3651 CCTTCTGCATGATCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 3710
| | | | |
QY 1261 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTTG 1301
| | | | |
Db 3711 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTTG 3751
| | | | |

RESULT 8

US-09-845-416-31
; Sequence 31, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: Del142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 31
; LENGTH: 4476
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-31

Query Match 100.0%; Score 1301; DB 13; Length 4476;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCCAGCAGTTTCAAGACGAAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
| | | | |
Db 2535 CGACTTTCCAGCAGTTTCAAGACGAAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 2594
| | | | |
QY 61 AACTAAAGAACCTGTAAATCATGAGTACTTGTGAGACTGTACGAAATATTTCTGACAGAGCA 120
| | | | |
Db 2595 AACTAAAGAACCTGTAAATCATGAGTACTTGTGAGACTGTACGAAATATTTCTGACAGAGCA 2654
| | | | |
QY 121 GCCTTTGGAAGGACTAGAGAAACTCTACAGGAGCCCGAGAGAGTGCCTCCTGAGGAGAG 180
| | | | |
Db 2655 GCCTTTGGAAGGACTAGAGAAACTCTACAGGAGCCCGAGAGAGTGCCTCCTGAGGAGAG 2714
| | | | |
QY 181 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 240
| | | | |
Db 2715 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 2774
| | | | |
QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT 300
| | | | |

Db 2775 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 2834
| | | | |
QY 301 CCAGGAACTTCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 360
| | | | |
Db 2835 CCAGGAACTTCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 2894
| | | | |
QY 361 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAAGATCACCTCGA 420
| | | | |
Db 2895 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAAGATCACCTCGA 2954
| | | | |
QY 421 GAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCCTCTGAAAGAGAACTGAGCCACGTCAG 480
| | | | |
Db 2955 GAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCCTCTGAAAGAGAACTGAGCCACGTCAG 3014
| | | | |
QY 481 TGACCTTGTCTCGCCAGCTTACCACCTTTGGGCATTTGAGCTCTCACCGTATAACCTCAGCAC 540
| | | | |
Db 3015 TGACCTTGTCTCGCCAGCTTACCACCTTTGGGCATTTGAGCTCTCACCGTATAACCTCAGCAC 3074
| | | | |
QY 541 TCTGGAAGACCTGAAACACCAAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 600
| | | | |
Db 3075 TCTGGAAGACCTGAAACACCAAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 3134
| | | | |
QY 601 GCAGCTGCATGAAGCCCAACAGGACCTTGGTCCAGCATCTCAGCACTTCTTCCACGTC 660
| | | | |
Db 3135 GCAGCTGCATGAAGCCCAACAGGACCTTGGTCCAGCATCTCAGCACTTCTTCCACGTC 3194
| | | | |
QY 661 TGTCCAGGTCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
| | | | |
Db 3195 TGTCCAGGTCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 3254
| | | | |
QY 721 CGAGACTCAAAACAACCTTGTGGGACCATCCAAATGACAGAGCTCTACAGTCTTTAGC 780
| | | | |
Db 3255 CGAGACTCAAAACAACCTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC 3314
| | | | |
QY 781 TGACCTGAATAATGTGAGTTCTCAGTTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
| | | | |
Db 3315 TGACCTGAATAATGTGAGTTCTCAGTTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3374
| | | | |
QY 841 GAAGGCCCTTCTGTTGATCTCTTGAGCCTGTGAGCTGCATGTGATGCTTGGACCAGCA 900
| | | | |
Db 3375 GAAGGCCCTTCTGTTGATCTCTTGAGCCTGTGAGCTGCATGTGATGCTTGGACCAGCA 3434
| | | | |
QY 901 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC 960
| | | | |
Db 3435 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC 3494
| | | | |
QY 961 TATTTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGGA 1020
| | | | |
Db 3495 TATTTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGGA 3554
| | | | |
QY 1021 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTTG 1080
| | | | |
Db 3555 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTTG 3614
| | | | |
QY 1081 CCTGTCTTTTAAAACTGGCATCATTTCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
| | | | |
Db 3615 CCTGTCTTTTAAAACTGGCATCATTTCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 3674
| | | | |
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCT 1200
| | | | |
Db 3675 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCT 3734
| | | | |
QY 1201 CCTTCTGCATGATCTATCCAAATTCGAGGACAGTGGGTGAAGTGCATCCTTTGGGGG 1260
| | | | |
Db 3735 CCTTCTGCATGATCTATCCAAATTCGAGGACAGTGGGTGAAGTGCATCCTTTGGGGG 3794
| | | | |
QY 1261 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTTG 1301
| | | | |
Db 3795 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTTG 3835
| | | | |

Db 3004 GCCTTTGGAAGGACTAGAGAAACTCTACCGAGAGCCAGAGAGCTGCCTCCTGAGGAGAG 3063
QY 181 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 240
Db 3064 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 3123
QY 241 AAAATTGAACCTGCACTCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 300
Db 3124 AAAATTGAACCTGCACTCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 3183
QY 301 CCAGGAACCTTCAAGAGCCACGGATGAGCTGGACCTCAAGCTCGCCAAAGCTGAGGTGAT 360
Db 3184 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCCAAAGCTGAGGTGAT 3243
QY 361 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCAAGATCACCTCGA 420
Db 3244 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCAAGATCACCTCGA 3303
QY 421 GAAAGTCAAGGCACCTTCGAGGAGAAAAATTGGCCTCTGAAAGAGAACTTGAGCCACGTCAA 480
Db 3304 GAAAGTCAAGGCACCTTCGAGGAGAAAAATTGGCCTCTGAAAGAGAACTTGAGCCACGTCAA 3363
QY 481 TGACCTTGCTCGCAGCTTACCACCTTTGGCACTTCCAGCTCTCACCGTATTAACCTCAGCAC 540
Db 3364 TGACCTTGCTCGCAGCTTACCACCTTTGGCACTTCCAGCTCTCACCGTATTAACCTCAGCAC 3423
QY 541 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAG 600
Db 3424 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAG 3483
QY 601 GCAGCTGCATGAAGCCACACAGGGACTTTTGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
Db 3484 GCAGCTGCATGAAGCCACACAGGGACTTTTGTCCAGCATCTCAGCACTTTCTTTCCACGTC 3543
QY 661 TGTCCAGGGTCCCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
Db 3544 TGTCCAGGGTCCCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 3603
QY 721 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTTAGC 780
Db 3604 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTTAGC 3663
QY 781 TGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
Db 3664 TGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3723
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCCCTTGGACCA 900
Db 3724 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCCCTTGGACCA 3783
QY 901 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCCTGCAGATTATTAATTGTTGACCC 960
Db 3784 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCCTGCAGATTATTAATTGTTGACCC 3843
QY 961 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCCTCTCTGCGTGA 1020
Db 3844 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCCTCTCTGCGTGA 3903
QY 1021 TATGTGTCTGAACCTGGTGTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGT 1080
Db 3904 TATGTGTCTGAACCTGGTGTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGT 3963
QY 1081 CCTGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAG 1140
Db 3964 CCTGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAG 4023
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCGAGCGAGCTGGGCCT 1200
Db 4024 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCGAGCGAGCTGGGCCT 4083
QY 1201 CCTTCTGCATGATTCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGTCATCCTTTGGGG 1260

Db 4084 CTTTCTGCATGATTCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 4143
QY 1261 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTTG 1301
Db 4144 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTTG 4184
RESULT 11
US-09-845-416-35
; Sequence 35, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 4848
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-35
Query Match 100.0%; Score 1301; DB 13; Length 4848;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
Db 2907 CGACTTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 2966
QY 61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db 2967 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 3026
QY 121 GCCTTTGGAAGGACTAGAGAAACTCTACAGGAGCCAGAGAGTGCCTCCTGAGGAGAG 180
Db 3027 GCCTTTGGAAGGACTAGAGAAACTCTACAGGAGCCAGAGAGTGCCTCCTGAGGAGAG 3086
QY 181 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 240
Db 3087 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 3146
QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 300
Db 3147 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 3206
QY 301 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 360
Db 3207 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 3266
QY 361 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAAGATCACCTCGA 420
Db 3267 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAAGATCACCTCGA 3326
QY 421 GAAAGTCAAGGCACCTTCGAGGAGAAAAATTGGCCTCTGAAAGAGAACGCTGAGCCACGTCAA 480
Db 3327 GAAAGTCAAGGCACCTTCGAGGAGAAAAATTGGCCTCTGAAAGAGAACGCTGAGCCACGTCAA 3386
QY 481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTGAGCTCTCACCGTATTAACCTCAGCAC 540
Db 3387 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTGAGCTCTCACCGTATTAACCTCAGCAC 3446
QY 541 TCTGGAAGACCTTGAAACACCAAGATGGAAGCTTCTGAGGTGGCCGTGAGGACCCGAGTCAG 600
Db 3447 TCTGGAAGACCTTGAAACACCAAGATGGAAGCTTCTGAGGTGGCCGTGAGGACCCGAGTCAG 3506
QY 601 GCAGCTGCATGAGCCCAACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTC 660

|||||
Db 3507 GCAGCTGCATGAAGCCACACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 3566
QY 661 TGTCCAGGTCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTACTATATCAACCA 720
Db 3567 TGTCCAGGTCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTACTATATCAACCA 3626
QY 721 CGAGACTCAACAACACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGC 780
Db 3627 CGAGACTCAACAACACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGC 3686
QY 781 TGACCTGAATAATGTCTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
Db 3687 TGACCTGAATAATGTCTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3746
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCTCAGCTGCATGTGATGCCCTTGGACCAAGCA 900
Db 3747 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCTCAGCTGCATGTGATGCCCTTGGACCAAGCA 3806
QY 901 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATGTTTGACCAC 960
Db 3807 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATGTTTGACCAC 3866
QY 961 TATTTATGACCGCTGGAGCAAGAGCAACAACAAATTTGGTCAACGTCCTCTCTGCGTGGGA 1020
Db 3867 TATTTATGACCGCTGGAGCAAGAGCAACAACAAATTTGGTCAACGTCCTCTCTGCGTGGGA 3926
QY 1021 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGCTGT 1080
Db 3927 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGCTGT 3986
QY 1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAG 1140
Db 3987 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAG 4046
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGGCT 1200
Db 4047 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGGCT 4106
QY 1201 CCTTCTGCATGATTTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 1260
Db 4107 CCTTCTGCATGATTTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 4166
QY 1261 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTTG 1301
Db 4167 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAATTTG 4207

RESULT 12
US-09-845-416-28
; Sequence 28, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 2001-04-30
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 4966
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-28
Query Match 100.0%; Score 1301; DB 13; Length 4966;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA 60
Db 3025 CGACTTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA 3084
QY 61 AACTAAAGAACCTGTAAATCATAGTACTCTTTGAGACTGTGACGAATATTTCTTGACAGAGCA 120
Db 3085 AACTAAAGAACCTGTAAATCATAGTACTCTTTGAGACTGTGACGAATATTTCTTGACAGAGCA 3144
QY 121 GCCTTTTGAAGGACTAGAGAAACTCTACCAAGGAGCCAGAGAGCTGCCCTCCTGAGGAGAG 180
Db 3145 GCCTTTTGAAGGACTAGAGAAACTCTACCAAGGAGCCAGAGAGCTGCCCTCCTGAGGAGAG 3204
QY 181 AGCCCAAGTGTCACTCCGCTTCTACGAAAGCAGGCTGAGAGGTTCAATACTGAGTGGGA 240
Db 3205 AGCCCAAGTGTCACTCCGCTTCTACGAAAGCAGGCTGAGAGGTTCAATACTGAGTGGGA 3264
QY 241 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT 300
Db 3265 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT 3324
QY 301 CCAGGAACCTCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 360
Db 3325 CCAGGAACCTCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 3384
QY 361 CAAGGGATCCTGGCAGCCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 420
Db 3385 CAAGGGATCCTGGCAGCCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA 3444
QY 421 GAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAAAGTGGCCACGTCAC 480
Db 3445 GAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAAAGTGGCCACGTCAC 3504
QY 481 TGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCAC 540
Db 3505 TGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCAC 3564
QY 541 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCTAG 600
Db 3565 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCTAG 3624
QY 601 GCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
Db 3625 GCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 3684
QY 661 TGTCCAGGTCCTTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTACTATATCAACCA 720
Db 3685 TGTCCAGGTCCTTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTACTATATCAACCA 3744
QY 721 CGAGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGTCTACCAAGTCTTTAGC 780
Db 3745 CGAGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGTCTACCAAGTCTTTAGC 3804
QY 781 TGACCTGAATAATGTCTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
Db 3805 TGACCTGAATAATGTCTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3864
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCTCAGCTGCATGTGATGCCCTTGGACCAAGCA 900
Db 3865 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCTCAGCTGCATGTGATGCCCTTGGACCAAGCA 3924
QY 901 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATGTTTGGACCAC 960
Db 3925 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATGTTTGGACCAC 3984
QY 961 TATTTATGACCGCTGGAGCAAGAGCAACAACAAATTTGGTCAACGTCCTCTCTGCGTGGGA 1020
Db 3985 TATTTATGACCGCTGGAGCAAGAGCAACAACAAATTTGGTCAACGTCCTCTCTGCGTGGGA 4044
QY 1021 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGCTGT 1080
Db 4045 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGCTGT 4104
QY 1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAG 1140

Db 4105 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTGGAAGACAAAGTACAG 4164
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACGAGCGCAGGCTGGGCT 1200
Db 4165 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACGAGCGCAGGCTGGGCT 4224
QY 1201 CCTTCTGCATGATCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1260
Db 4225 CCTTCTGCATGATCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 4284
QY 1261 CAGTAACATTGAGCCAAAGTGTCCGAGCTGCTTCCAATTTG 1301
Db 4285 CAGTAACATTGAGCCAAAGTGTCCGAGCTGCTTCCAATTTG 4325

RESULT 13
US-09-845-416-34
; Sequence 34, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 4990
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-34

Query Match 100.0%; Score 1301; DB 13; Length 4990;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA 60
Db 3049 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA 3108
QY 61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGACCA 120
Db 3109 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGACCA 3168
QY 121 GCCTTTGGAAGGACTAGAGAAACTCTACAGGAGCCCAGAGAGTGCCTCCTGAGGAGAG 180
Db 3169 GCCTTTGGAAGGACTAGAGAAACTCTACAGGAGCCCAGAGAGTGCCTCCTGAGGAGAG 3228
QY 181 AGCCCAGAAATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 240
Db 3229 AGCCCAGAAATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 3288
QY 241 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 300
Db 3289 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 3348
QY 301 CCAGGAACCTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 360
Db 3349 CCAGGAACCTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 3408
QY 361 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 420
Db 3409 CAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 3468
QY 421 GAAAGTCAAGGCACCTTCAGGAGAGAAATTCGCCCTCTGAAAGAGAACGTTGAGCCACGTCAA 480
Db 3469 GAAAGTCAAGGCACCTTCAGGAGAGAAATTCGCCCTCTGAAAGAGAACGTTGAGCCACGTCAA 3528

QY 481 TGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTACAGCTCTCACCGTATAACCTCAGCAC 540
Db 3529 TGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTACAGCTCTCACCGTATAACCTCAGCAC 3588
QY 541 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCGTCGAGGACCGAGTCAG 600
Db 3589 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCGTCGAGGACCGAGTCAG 3648
QY 601 GCAGCTGCATGAAGCCACACAGGACTTTGGTCCAGCATCTCAGCACTCTCAGCACTTCTTTCCACGTC 660
Db 3649 GCAGCTGCATGAAGCCACACAGGACTTTGGTCCAGCATCTCAGCACTCTCAGCACTTCTTTCCACGTC 3708
QY 661 TGTCCAGGTCCTTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
Db 3709 TGTCCAGGTCCTTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 3768
QY 721 CGAGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
Db 3769 CGAGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 3828
QY 781 TGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
Db 3829 TGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3888
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCTTTGGACCAGCA 900
Db 3889 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCTTTGGACCAGCA 3948
QY 901 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC 960
Db 3949 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC 4008
QY 961 TATTTATGACCGCCTGGAGCAAGACCAACAATTTGGTCAACGTCCTCTCTCGCTGGA 1020
Db 4009 TATTTATGACCGCCTGGAGCAAGACCAACAATTTGGTCAACGTCCTCTCTCGCTGGA 4068
QY 1021 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACCAACAGGAGGATCCGTGT 1080
Db 4069 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACCAACAGGAGGATCCGTGT 4128
QY 1081 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
Db 4129 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 4188
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACGAGCGCAGGCTGGGCT 1200
Db 4189 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACGAGCGCAGGCTGGGCT 4248
QY 1201 CCTTCTGCATGATCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1260
Db 4249 CCTTCTGCATGATCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 4308
QY 1261 CAGTAACATTGAGCCAAAGTGTCCGAGCTGCTTCCAATTTG 1301
Db 4309 CAGTAACATTGAGCCAAAGTGTCCGAGCTGCTTCCAATTTG 4349

RESULT 14
US-09-845-416-36
; Sequence 36, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36

```

; LENGTH: 5060
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-36

Query Match      100.0%; Score 1301; DB 13; Length 5060;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTTCAGCAGTTTCAGAAAGCAGAACCGATGTACATAGGGCCCTCAAGAGGGAATTGAA 60
Db 3119 CGACTTTCAGCAGTTTCAGAAAGCAGAACCGATGTACATAGGGCCCTCAAGAGGGAATTGAA 3178

QY 61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAAATATTTCTGACAGAGCA 120
Db 3179 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAAATATTTCTGACAGAGCA 3238

QY 121 GCCTTTGGAAGGACTAGAGAAACTCTACAGGAGCCCGAGAGAGCTGCCTCCTGAGGAGAG 180
Db 3239 GCCTTTGGAAGGACTAGAGAAACTCTACAGGAGCCCGAGAGAGCTGCCTCCTGAGGAGAG 3298

QY 181 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 240
Db 3299 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 3358

QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 300
Db 3359 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 3418

QY 301 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCGCAAGCTGAGGTGAT 360
Db 3419 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCGCAAGCTGAGGTGAT 3478

QY 361 CAAGGGATCCTGGAGCCCGTGGCGATCTCTCATTGACTCTCTCCAAAGATCACCTCGA 420
Db 3479 CAAGGGATCCTGGAGCCCGTGGCGATCTCTCATTGACTCTCTCCAAAGATCACCTCGA 3538

QY 421 GAAAGTCAAGGCACCTTCAGAGAGAAATTGGGCTCTGAAAGAGAACCTGAGCCACGTCAA 480
Db 3539 GAAAGTCAAGGCACCTTCAGAGAGAAATTGGGCTCTGAAAGAGAACCTGAGCCACGTCAA 3598

QY 481 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCCGATTAACCTCAGCAC 540
Db 3599 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCCGATTAACCTCAGCAC 3658

QY 541 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 600
Db 3659 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 3718

QY 601 GCAGCTGCATGAAGCCCAACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTC 660
Db 3719 GCAGCTGCATGAAGCCCAACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTC 3778

QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
Db 3779 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 3838

QY 721 CGAGACTCAAACAACTTGCTGGGACCAATCCCAAATGACAGAGCTCTACCACTCTTTAGC 780
Db 3839 CGAGACTCAAACAACTTGCTGGGACCAATCCCAAATGACAGAGCTCTACCACTCTTTAGC 3898

QY 781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGACTGCCATGAACTCCGAAGACTGCA 840
Db 3899 TGACCTGAATAATGTCAGATTCTCAGCTTATAGACTGCCATGAACTCCGAAGACTGCA 3958

QY 841 GAAGGCCCTTTGCTTGGAATCTTTGAGCCCTGTGAGCTGCATGTGATGCTTGACCAGCA 900
Db 3959 GAAGGCCCTTTGCTTGGAATCTTTGAGCCCTGTGAGCTGCATGTGATGCTTGACCAGCA 4018

QY 901 CAACTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAAATGTTTGACCCAC 960
Db 4019 CAACTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAAATGTTTGACCCAC 4078
```

```

QY 961 TATTTATGACCGCTGGAGCAAGACACACAATAATTTGGTCAACGTCCCTCTCTGCGTGG 1020
Db 4079 TATTTATGACCGCTGGAGCAAGACACACAATAATTTGGTCAACGTCCCTCTCTGCGTGG 4138

QY 1021 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 1080
Db 4139 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 4198

QY 1081 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCAACATTTGGAAGACAAGTACAG 1140
Db 4199 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCAACATTTGGAAGACAAGTACAG 4258

QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 1200
Db 4259 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 4318

QY 1201 CCTTCTGCTATGATTTCTATCCAAATTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1260
Db 4319 CCTTCTGCTATGATTTCTATCCAAATTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 4378

QY 1261 CAGTAACATTCAGCCAAAGTGTCCGGAGCTGCTTCCAATTTG 1301
Db 4379 CAGTAACATTCAGCCAAAGTGTCCGGAGCTGCTTCCAATTTG 4419
```

```

RESULT 15
US-09-845-416-27
; Sequence 27, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 5149
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-27
```

```

Query Match      100.0%; Score 1301; DB 13; Length 5149;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTTCAGCAGTTTCAGAAAGCAGAACCGATGTACATAGGGCCCTCAAGAGGGAATTGAA 60
Db 3208 CGACTTTCAGCAGTTTCAGAAAGCAGAACCGATGTACATAGGGCCCTCAAGAGGGAATTGAA 3267

QY 61 AACTAAAGAACCTGTAATCATGAGTACTCTTTGAGACTGTACGAAATATTTCTGACAGAGCA 120
Db 3268 AACTAAAGAACCTGTAATCATGAGTACTCTTTGAGACTGTACGAAATATTTCTGACAGAGCA 3327

QY 121 GCCTTTGGAAGGACTAGAGAAACTCTACAGGAGCCCGAGAGAGCTGCCTCCTGAGGAGAG 180
Db 3328 GCCTTTGGAAGGACTAGAGAAACTCTACAGGAGCCCGAGAGAGCTGCCTCCTGAGGAGAG 3387

QY 181 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 240
Db 3388 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 3447

QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 300
Db 3448 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 3507

QY 301 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGAT 360
Db 3508 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGAT 3567
```

| | | | |
|----|------|--|------|
| QY | 361 | CAAGGGATCCTGGCAGCCCGTGGGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA | 420 |
| Db | 3568 | CAAGGGATCCTGGCAGCCCGTGGGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGA | 3627 |
| QY | 421 | GAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAA | 480 |
| Db | 3628 | GAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAA | 3687 |
| QY | 481 | TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAACCTCAGCAC | 540 |
| Db | 3688 | TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAACCTCAGCAC | 3747 |
| QY | 541 | TCTGGAAGACCTGAACACACCAGATGGAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAG | 600 |
| Db | 3748 | TCTGGAAGACCTGAACACACCAGATGGAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAG | 3807 |
| QY | 601 | GCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACATTTCTTTCCACGTC | 660 |
| Db | 3808 | GCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACATTTCTTTCCACGTC | 3867 |
| QY | 661 | TGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA | 720 |
| Db | 3868 | TGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA | 3927 |
| QY | 721 | CGAGACTCAAAACAACTTGTGTGGACCAATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC | 780 |
| Db | 3928 | CGAGACTCAAAACAACTTGTGTGGACCAATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC | 3987 |
| QY | 781 | TGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA | 840 |
| Db | 3988 | TGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA | 4047 |
| QY | 841 | GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGATGATGCCTTGGACCAGCA | 900 |
| Db | 4048 | GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGATGATGCCTTGGACCAGCA | 4107 |
| QY | 901 | CAACCTCAAGCAAAATGACCGCCCATGGATATCCTGCAGATTATTAATTGTTGACCAC | 960 |
| Db | 4108 | CAACCTCAAGCAAAATGACCGCCCATGGATATCCTGCAGATTATTAATTGTTGACCAC | 4167 |
| QY | 961 | TATTTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTCGGTGGA | 1020 |
| Db | 4168 | TATTTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTCGGTGGA | 4227 |
| QY | 1021 | TATGTGCTGAACGTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT | 1080 |
| Db | 4228 | TATGTGCTGAACGTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT | 4287 |
| QY | 1081 | CCTGTCTTTTAAACTGGCATCATTTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG | 1140 |
| Db | 4288 | CCTGTCTTTTAAACTGGCATCATTTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG | 4347 |
| QY | 1141 | ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT | 1200 |
| Db | 4348 | ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT | 4407 |
| QY | 1201 | CCTTCTGCATGATCTATCCAAATTCGAAGACAGTTGGGTGAGTTGCATCCTTTGGGGG | 1260 |
| Db | 4408 | CCTTCTGCATGATCTATCCAAATTCGAAGACAGTTGGGTGAGTTGCATCCTTTGGGGG | 4467 |
| QY | 1261 | CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTG | 1301 |
| Db | 4468 | CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTG | 4508 |

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 11:04:35 ; Search time 2439.51 Seconds
(without alignments)
12961.661 Million cell updates/sec

Title: US-09-845-416-10_COPY_1800_3100
Perfect score: 1301
Sequence: 1 cgactttccagcagttcaga.....ccggagctgcttccaatttg 1301

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

| | |
|-----|--------------|
| 1: | em_estba:* |
| 2: | em_esthum:* |
| 3: | em_estin:* |
| 4: | em_estmu:* |
| 5: | em_estov:* |
| 6: | em_estpl:* |
| 7: | em_estro:* |
| 8: | em_htc:* |
| 9: | gb_est1:* |
| 10: | gb_est2:* |
| 11: | gb_htc:* |
| 12: | gb_est3:* |
| 13: | gb_est4:* |
| 14: | gb_est5:* |
| 15: | em_estfun:* |
| 16: | em_estom:* |
| 17: | em_gss_hum:* |
| 18: | em_gss_inv:* |
| 19: | em_gss_pln:* |
| 20: | em_gss_vrt:* |
| 21: | em_gss_fun:* |
| 22: | em_gss_mam:* |
| 23: | em_gss_mus:* |
| 24: | em_gss_pro:* |
| 25: | em_gss_rod:* |
| 26: | em_gss_phg:* |
| 27: | em_gss_vrl:* |
| 28: | gb_gss1:* |
| 29: | gb_gss2:* |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 584 | 44.9 | 824 | 9 | AL556247 |
| 2 | 582.4 | 44.8 | 1098 | 13 | BX365572 |
| 3 | 573 | 44.0 | 620 | 13 | BQ640063 |
| 4 | 571.4 | 43.9 | 801 | 14 | CB991394 |

| | | | | | |
|----|-------|------|------|----|-----------|
| 5 | 537.8 | 41.3 | 770 | 10 | BG719710 |
| 6 | 535 | 41.1 | 3753 | 11 | AK081426 |
| 7 | 526.4 | 40.5 | 1298 | 11 | AK087829 |
| 8 | 526.4 | 40.5 | 2135 | 11 | AK013510 |
| 9 | 526.4 | 40.5 | 4437 | 11 | AK036936 |
| 10 | 524.8 | 40.3 | 1384 | 11 | AK075809 |
| 11 | 511 | 39.3 | 797 | 14 | CB960722 |
| 12 | 502.8 | 38.6 | 652 | 10 | BB629984 |
| 13 | 491.6 | 37.8 | 704 | 10 | BB610411 |
| 14 | 479.6 | 36.9 | 854 | 9 | AI196693 |
| 15 | 470.8 | 36.2 | 717 | 14 | CB527785 |
| 16 | 455.8 | 35.0 | 591 | 2 | HSM083521 |
| 17 | 439 | 33.7 | 578 | 2 | HSM075761 |
| 18 | 436.6 | 33.6 | 665 | 14 | BY742604 |
| 19 | 431.4 | 33.2 | 898 | 10 | BF182065 |
| 20 | 427.8 | 32.9 | 843 | 14 | CA988247 |
| 21 | 426.6 | 32.8 | 3051 | 11 | BC036095 |
| 22 | 410.4 | 31.5 | 599 | 10 | BB666688 |
| 23 | 407.2 | 31.3 | 554 | 14 | CB613696 |
| 24 | 404.4 | 31.1 | 493 | 14 | CA888041 |
| 25 | 397.4 | 30.5 | 495 | 14 | CA894775 |
| 26 | 392.2 | 30.1 | 650 | 14 | BY714491 |
| 27 | 389 | 29.9 | 794 | 13 | BU424348 |
| 28 | 385.2 | 29.6 | 684 | 9 | AL641565 |
| 29 | 383.4 | 29.5 | 483 | 14 | CA893902 |
| 30 | 381.8 | 29.3 | 423 | 9 | AA460476 |
| 31 | 380.4 | 29.2 | 633 | 9 | AL796733 |
| 32 | 375.4 | 28.9 | 514 | 9 | AL871560 |
| 33 | 365.4 | 28.1 | 655 | 14 | CA558919 |
| 34 | 349.6 | 26.9 | 546 | 4 | BX516597 |
| 35 | 347.2 | 26.7 | 665 | 10 | BE201973 |
| 36 | 337.6 | 25.9 | 541 | 9 | AL894729 |
| 37 | 330 | 25.4 | 608 | 9 | AL672616 |
| 38 | 329 | 25.3 | 532 | 10 | BE334408 |
| 39 | 325.6 | 25.0 | 532 | 12 | BJ036615 |
| 40 | 325.6 | 25.0 | 883 | 14 | CD557811 |
| 41 | 323.4 | 24.9 | 644 | 13 | BU313510 |
| 42 | 320.6 | 24.6 | 406 | 14 | CA871483 |
| 43 | 315.4 | 24.2 | 541 | 12 | BI546771 |
| 44 | 313 | 24.1 | 656 | 9 | AL869539 |
| 45 | 310 | 23.8 | 488 | 12 | BM488542 |

ALIGNMENTS

RESULT 1
AL556247
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AL556247 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
cDNA clone CSODK001YB17 5-PRIME, mRNA sequence.
AL556247 GI:31278051
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 824)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12898746.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1955.r
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSODK001CA09QP1.
Location/Qualifiers

FEATURES

```
source
1. .824
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK001YB17"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      211 a      210 c      206 g      197 t
ORIGIN
      44.9%;   Score 584;   DB 9;   Length 824;
      100.0%;   Pred. No. 7.1e-160;
      0;   Mismatches 0;   Indels 0;   Gaps 0;
718 CCACGAGACTCAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTT 777
195 CCACGAGACTCAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTT 254
778 AGCTGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 837
255 AGCTGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 314
838 GCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTGGACCA 897
315 GCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTGGACCA 374
898 GCACAAACCTCAAGCAAAATGACCCAGCCCATGGATATCCCTGACAGATTATTAATTGTTGAC 957
375 GCACAAACCTCAAGCAAAATGACCCAGCCCATGGATATCCCTGACAGATTATTAATTGTTGAC 434
958 CACTATTTATGACCGCCTGGAGCGAGAGACACAACAAATTTGGTCAACGCTCCCTCTCTGCGT 1017
435 CACTATTTATGACCGCCTGGAGCGAGAGACACAACAAATTTGGTCAACGCTCCCTCTCTGCGT 494
1018 GGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 1077
495 GGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 554
1078 TGTCTCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAGTA 1137
555 TGTCTCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAGTA 614
1138 CAGATACCTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCGGAGGCTGGG 1197
615 CAGATACCTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCGGAGGCTGGG 674
1198 CCTCCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGTGAAGTTGCATCCTTTGG 1257
675 CCTCCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGTGAAGTTGCATCCTTTGG 734
1258 GGGCAGTAACATTGAGCCAAAGTGTCGGAGCTGCTTCCAAATTG 1301
735 GGGCAGTAACATTGAGCCAAAGTGTCGGAGCTGCTTCCAAATTG 778

RESULT 2
BX365572
LOCUS
DEFINITION BX365572 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DK001YB21 5-PRIME, mRNA sequence.
ACCESSION BX365572
VERSION BX365572.1 GI:30366927
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1098)
```

```
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1955.r
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS1AK001ZB11QP1.
FEATURES
Location/Qualifiers
1. .1098
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK001YB21"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      286 a      271 c      268 g      262 t      11 others
ORIGIN
      44.8%;   Score 582.4;   DB 13;   Length 1098;
      99.8%;   Pred. No. 2.5e-159;
      0;   Mismatches 1;   Indels 0;   Gaps 0;
Query Match
Best Local Similarity
Matches 583; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
718 CCACGAGACTCAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTT 777
204 CCACGAGACTCAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTT 263
778 AGTGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 837
264 AGTGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 323
838 GCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTGGACCA 897
324 GCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTGGACCA 383
898 GCACAACTCAAGCAAAATGACCCAGCCCATGGATATCCCTGACAGATTATTAATTGTTGAC 957
384 GCACAACTCAAGCAAAATGACCCAGCCCATGGATATCCCTGACAGATTATTAATTGTTGAC 443
958 CACTATTTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGCTCCCTCTCTGCGT 1017
444 CACTATTTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGCTCCCTCTCTGCGT 503
1018 GGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 1077
504 GGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 563
1078 TGTCTCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAGTA 1137
564 TGTCTCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAGTA 623
1138 CAGATACCTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCGGAGGCTGGG 1197
624 CAGATACCTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCGGAGGCTGGG 683
1198 CCTCCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGTGAAGTTGCATCCTTTGG 1257
684 CCTCCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGTGAAGTTGCATCCTTTGG 743
1258 GGGCAGTAACATTGAGCCAAAGTGTCGGAGCTGCTTCCAAATTG 1301
744 GGGCAGTAACATTGAGCCAAAGTGTCGGAGCTGCTTCCAAATTG 787
```

```
RESULT 3
30640063
LOCUS
DEFINITION
he23g04.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he
Homo sapiens cDNA clone he23g04 5', mRNA sequence.
BQ640063
ACCESSION
BQ640063.1 GI:21764522
VERSION
EST.
KEYWORDS
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 620)
Wistow,G., Bernstein,S.L., Wyatt,M.K., Ray,S., Behal,A., Touchman
,J.W., Bouffard,G., Smith,D. and Peterson,K.
Expressed sequence tag analysis of human retina for the NEIBank
Project: Retbindin, an abundant, novel retinal cDNA and alternative
splicing of other retina-preferred gene transcripts
Mol. Vis. 8 (4), (2002) In press
Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 23 row: g column: 04
Seq primer: M13RP1 reverse primer (ABI).
Location/Qualifiers
1..620
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="he23g04"
/tissue_type="Retina"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Human Retina cDNA (Un-normalized, unamplified
): hd/he"
/note="Organ: Eye; Vector: pSPORT1; Neural retina tissue
was dissected from two 80 year old donors with no observed
eye disease. 100ug of total RNA was used for library
construction. A directionally cloned cDNA library in the
pSPORT1 vector (Life Technologies) was constructed at
Bioserve Biotechnology (Laurel MD) essentially following
the protocols of the Superscript Plasmid System full
details of which are contained in the manufacturer's
instruction manual (http://www.lifetech.com/). First
strand synthesis was carried out using a Not I
primer-adaptor [5'-pGACTAGTTCAGATCGGAGCGGCCCT(T)15-3'
]. EST analysis was performed on the unamplified library
at the NIH Intramural Sequencing Center (NISC)."
```

```
BASE COUNT 165 a 148 c 144 g 163 t
ORIGIN
Query Match 44.0%; Score 573; DB 13; Length 620;
Best Local Similarity 100.0%; Pred. No. 9.9e-157;
Matches 573; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
2Y 729 AAACAACCTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGA 788
|
|
|
Db 1 AAACAACCTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGA 60
|
|
|
2Y 789 ATAATGTCAGATTCTCAGCTTTATAGGACTGCCATGAACTCCGAAGACTGCAGAAGGCC 848
|
|
|
Db 61 ATAATGTCAGATTCTCAGCTTTATAGGACTGCCATGAACTCCGAAGACTGCAGAAGGCC 120
|
|
|
2Y 849 TTTGCTTGATCTCTTGAGCCCTGTCAGCTGTCATGTGATGCTTGACCGACGACACACCTCA 908
|
|
|
Db 121 TTTGCTTGATCTCTTGAGCCCTGTCAGCTGTCATGTGATGCTTGACCGACGACACACCTCA 180
|
|
|
2Y 909 AGCAAAATGACGACCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACATTATG 968
|
|
|
```

```
Db 181 AGCAAAATGACGACCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACATTATG 240
QY 969 ACCGCCTGGAGCAAGAGACAAACAATTTGGTCAACGTCCTCTCTCTCGGTGGATATGTGTC 1028
|
|
|
Db 241 ACCGCCTGGAGCAAGAGACAAACAATTTGGTCAACGTCCTCTCTCTCGGTGGATATGTGTC 300
|
|
|
QY 1029 TGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGCTTT 1088
|
|
|
Db 301 TGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGCTTT 360
|
|
|
QY 1089 TTAAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAGTACAGATACTTTT 1148
|
|
|
Db 361 TTAAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAGTACAGATACTTTT 420
|
|
|
QY 1149 TCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCCTCCTTCTGC 1208
|
|
|
Db 421 TCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCCTCCTTCTGC 480
|
|
|
QY 1209 ATGATTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGCAAGTAACA 1268
|
|
|
Db 481 ATGATTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGCAAGTAACA 540
|
|
|
QY 1269 TTGAGCCAAGTGTCCGGAGTGTCTTCCAATTG 1301
|
|
|
Db 541 TTGAGCCAAGTGTCCGGAGTGTCTTCCAATTG 573
```

```
RESULT 4
CB991394 801 bp mRNA linear EST 01-MAY-2003
AGENCOURT 13627932 NIH_MGC_148 Homo sapiens cDNA clone
IMAGE:30336570 5', mRNA sequence.
CB991394
CB991394.1 GI:30285818
EST.
Homo sapiens (human)
Homo sapiens
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 801)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM360 row: d column: 19
High quality sequence stop: 621.
```

```
Location/Qualifiers
1..801
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30336570"
/tissue_type="pre-eclamptic placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH MGC 148"
/note="Organ: placenta; Vector: pBluescriptR; Site 1:
ali-XhoI; Site 2: BamH; Library is oligo-dT primed and
directionally cloned using primer
5'-TTTTTTTTTTTTTTTN-3', size-selected for average insert
size 2.3 kb and normalized to ROT 5. This is a primary
library enriched for full-length clones and constructed
using the Cap-trapper method (Carninci, in preparation).
Library constructed by M. Brownstein (NIH/NHGRI,
National Institutes of Health). Note: this is a NIH_MGC
```


Library." 204 a 210 c 199 g 188 t

Query Match 43.9%; Score 571.4; DB 14; Length 801;
Best Local Similarity 99.7%; Pred. No. 3.5e-156;
Matches 583; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 718 CCACGAGACTCAAACAACCTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTT 777
Db 140 CCACGAGACTCAAACAACCTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTT 199

QY 778 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGACTGCCATGAAACTCCGAAGACT 837
Db 200 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGACTGCCATGAAACTCCGAAGACT 259

QY 838 GCAGAGGCGCCCTTTGCTTGGATCTCTTGAGCCCTGTACGCTGCATGTGATGCCCTGGACCA 897
Db 260 GCAGAGGCGCCCTTTGCTTGGATCTCTTGAGCCCTGTACGCTGCATGTGATGCCCTGGACCA 319

QY 898 GCACAACTCAAGCAAAATGACAGCCCATGGATATCTGCAGATTATTAAATTGTTGAC 957
Db 320 GCACAACTCAAGCAAAATGACAGCCCATGGATATCTGCAGATTATTAAATTGTTGAC 379

QY 958 CACTATTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGT 1017
Db 380 CACTATTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGT 439

QY 1018 GGATATGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 1077
Db 440 GGATATGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 499

QY 1078 TGTCTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAGTA 1137
Db 500 TGTCTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAGTA 559

QY 1138 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGGCTGGG 1197
Db 560 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGGCTGGG 619

QY 1198 CCTCTCTCTGCATGATTCTATCCAAATTTCCAAAGACAGTTGGTGAAGTTGCATCCTTTGG 1257
Db 620 CCTCTCTCTGCATGATTCTATCCAAATTTCCAAAGACAGTTGGTGAAGTTGCATCCTTTGG 679

QY 1258 GGGCAGTAACATTGAGCCCAAGTGTC-GGAGCTGCTTCCAATTG 1301
Db 680 GGGCAGTAACATTGAGCCCAAGTGTCCTGGGAGCTGCTTCCAATTG 724

RESULT 5
BG719710
LOCUS
DEFINITION 602690430F1 NIH_MGC_97 770 bp mRNA linear EST 08-MAY-2001
mRNA sequence.
ACCESSION BG719710
VERSION BG719710.1 GI:13998897
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 770)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10731 row: 1 column: 08
High quality sequence stop: 767.

FEATURES
source
Location/Qualifiers
1..770
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4822807"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: Sali-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 189 a 199 c 198 g 184 t
ORIGIN

Query Match 41.3%; Score 537.8; DB 10; Length 770;
Best Local Similarity 97.4%; Pred. No. 2.5e-146;
Matches 568; Conservative 0; Mismatches 12; Indels 3; Gaps 2;

QY 718 CCACGAGACTCAAACAACCTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTT 777
Db 150 CCACGAGACTCAAACAACCTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTT 209

QY 778 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGACTGCCATGAAACTCCGAAGACT 837
Db 210 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGACTGCCATGAAACTCCGAAGACT 269

QY 838 GCAGAGGCGCCCTTTGCTTGGATCTCTTGAGCCCTGTACGCTGCATGTGATGCCCTGGACCA 897
Db 270 GCAGAGGCGCCCTTTGCTTGGATCTCTTGAGCCCTGTACGCTGCATGTGATGCCCTGGACCA 329

QY 898 GCACAA - CCTCAAGCAAAATGACACGCCCCATGGATATCTGCAGATTATTAAATTGTTG 955
Db 330 GCACAAATCCTCAAGTCAGATGACACGCCCCATGGATATCTGCAGATTATTAAATTGTTG 389

QY 956 ACCACTATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGTC 1015
Db 390 ACCACTATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGTC 449

QY 1016 GTGGATATGTGTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATC 1075
Db 450 GTGGATATGTGTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATC 509

QY 1076 CGTGTCTCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTGGAAGACAAG 1135
Db 510 CGTGTCTCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTGGAAGACAAG 569

QY 1136 TACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCGAGGCTG 1195
Db 570 TACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCGAGGCTG 628

QY 1196 GGCCTCCTCTGTCATGATTCTATCCAAATTTCCAAAGACAGTTGGTGAAGTTGCATCCTTT 1255
Db 629 GGCCTCCTCTGTCATGATTCTATCCAAATTTCCAAAGACAGTGGGTGAAGTTGCATCCTTT 688

QY 1256 GGGGGCAGTAACATTGAGCCCAAGTGTCCGAGCTGCTTCCAAT 1298
Db 689 GGGGGCAGTAACATAGAGCCAAGTGTCCGAGCTGCTTCCAAT 731

RESULT 6
AK081426
LOCUS
DEFINITION Mus musculus 16 days embryo head cdna, RIKEN full-length enriched

QY 74 GTAATCATGAGTACTCTTTGAGACTGTACGAATATTTTGTGACAGAGCAGCCCTTTGGAAGGA 133
Db 652 TACATCTCCTCTGTGCTTGAATCGCCAGGCTTTTGTGCCAGCATCCATTGAAGAA 711
QY 134 CTAGAGAAACTCTACCAAGGAGCCAGAGAGCTGCCTCTGAGGAGAGAGCCAGAAATGTC 193
Db 712 TTAGAGAGTCTCAATCTGAGAGTAAAGATACCTCCCCAGACAGCGGATTCAAAACCTT 771
QY 194 ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATFACTGAGTGGGAAAAATTGAACCTG 253
Db 772 AGTCGCTTTGTATGGAAGCAGGCAACAGTGGCCAGTGAAGTGTGGGAGAAGCTGACAGCC 831
QY 254 CACTCCGCTGACTGGCAGAGAGAAATAATAGATGAGACCTTTGAAAAGACTCCAGGAACTTCAA 313
Db 832 CGCTGTGTAGATCAGCATCGCCACATTTGAGCATACTCTGGAACATCTATTGGAGATCCAA 891
QY 314 GAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGG 373
Db 892 GGGGCAATGGAGGAATTGAGCAGTACTTTGACCCCAAGCAGAGGGAGTCCGAGCCAGCTGG 951
QY 374 CAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCA 433
Db 952 GAGCCCATAGGAGATCTCTTTATTGATTCCTCCAGAGCATATCCAAGCCATCAAGCTA 1011
QY 434 CTTGAGGAGAAATTCGCCCTCTGAAAGAGAACGTGAGCCAGCTCAATGACCTTGCTCGC 493
Db 1012 TTCAAAGAAGAAATCTCTCCTGTGAAAGATGGGTGAAGTTAGTGAATGATCTGGCCAC 1071
QY 494 CAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATACCTCAAGCTCAGCACTCTGGAAGACTG 553
Db 1072 CAGCTTGCCATCTCTGATGTGCACCTTGTCAATGGAGAAATTCAGGGCTCTGGAACAGATC 1131
QY 554 AACACCAGATGGAAGCTTCTGCAAGTGGCGTTCGAGGACCGAGTCAAGCAGCTGCATGAA 613
Db 1132 AACATCCGGTGGAAACAGCTCCAGGTGTGAGTGGCTGAGAGGCTTAAGCAACTCCAGGAT 1191
QY 614 GCCCAGAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTGTCCAGGGTCCC 673
Db 1192 GCCCAGGAGCTTTGGCCCTGGGTCAAGCACTTCTCTCCACTTCTGTCCAAGTTCCC 1251
QY 674 TGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACACAGAGACTCAAAACA 733
Db 1252 TGGGAAAGAGCATTTTCTCCCAATAAAGTCCCTACTACATCAACACAGGCTCAGACC 1311
QY 734 ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAAT 793
Db 1312 ACATGCTGGGACCATCCTAAGATGACTGAGTATATACCAACCTTAGCCGATCTGAACAC 1371
QY 794 GTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGAGAGGCCCTTTGC 853
Db 1372 ATTAAGTTCTCAGCTTACCGCACTGCCATGAAGCTCCGAGAGTCCAGAAGGCCCTGGC 1431
QY 854 TTGGATCTCTTGAGCCTGTGACGTGCATGTGATGCTTGGACCGACGACCAACCTCAAGCAA 913
Db 1432 CTGGATCTGGTAACCTTTAACTACAGCTCTGGAGATCTTCAATGAGCATGACTTGCAGGCC 1491
QY 914 AATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTATGACCGC 973
Db 1492 AGTGAACATGTGATGGATGTGGTGGAGGTCAATCACTGCTTGACTGCTTGTATGAACGA 1551
QY 974 CTGGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTGCGTGGATATGTGTGAAC 1033
Db 1552 CTGGAGGAGGAAAGAGGCATCTGGTCAATGTGCACTGTGTGTAGACATGAGCCTCAAC 1611
QY 1034 TGGCTGCTGAATGTTTATGATACGGGACGAGGAGGATCCGTCCTGCTCTTTTAA 1093
Db 1612 TGCTCCTCAATGTTTGTATAGTGGTGGAGTGGGAAAGATGCGAGCATTTGCTTTAAG 1671
QY 1094 ACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAAGACAAGTACAGATACCTTTTCAAG 1153
Db 1672 ACTGGCATCGCATGCCTGTGTGGCAGCGAAGTGAAGAAAGAACTTCAGTATCTCTTCAGC 1731
QY 1154 CAAGTGGCAAGTTCAACAGGATTTTGTGACCGAGGCTGGGCCCTCTCTGCAATGAT 1213

Db 1732 CAAGTAGCCAATTCAGGAGCCAGTGTGATCAAGCCCATCTCGGTGCCCTGCTTCATGAA 1791
QY 1214 TCTATCCAAATTCCAAGACAGTTGGTGAAGTTGATCCTTTTGGGGGCGAGTAACATTGAG 1273
Db 1792 GCCATTCAAGTGCCTCCGTCAGCTGGTGAAGTGGCAGCATTGGGGGCGAGCAATGTGGAG 1851
QY 1274 CCAAGTGTCCGGAGCTGCTTCCAATTT 1300
Db 1852 CCAAGTGTCCGTAAGTGTGCTTTCGCTTT 1878
RESULT 7
AK087829 1298 bp mRNA linear HTC 05-DEC-2002
LOCUS Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN
DEFINITION full-length enriched library, clone:E330026B12 product:dystrophin,
muscular dystrophy, full insert sequence.
ACCESSION AK087829 GI:26352777
VERSION AK087829.1
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
AUTHORS Carninci, P., and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
2
REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
3
REFERENCE
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
4
REFERENCE
AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Sakamoto, N., Sasaki, H.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
and Hayashizaki, Y.
TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409 (6821), 685-690 (2001)
MEDLINE 21085660
PUBMED 11217851
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 1298)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@sc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/
FEATURES Location/Qualifiers
source 1..1298
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:E330026B12"
/db_xref="taxon:10090"
/clone="E330026B12"
/sex="female"
/tissue_type="ovary"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days pregnant adult"
120..1142
/note="unnamed protein product; dystrophin, muscular dystrophy (MGD|GI:94909, GB|NM_007868, evidence: BLASTN, 100%, match=999)
putative"
/codon_start=1
/protein_id="BAC40019.1"
/db_xref="GI:26352778"
/translation="MREHLKGHETQTTCWDHPKMTELYQSADLNNVRFSAVRTAMKL RRLQKALCLDLSLSAACDLDQNLKQNDQPMILQINCLTIYDRLEQEHNNLVN VPLCDVMCLNWLNNVDTGRIRVLSFKTGIIISLKAHLEDKYRYLFQVASSTGF CDQRRLGLLLHDSIQIPRLGVEVAFSGSNIEPSVRSFQFANNKPEIEAALFLDWMR LEPQSMVWLPVLHVRVAAAEATKHQAACNICECPILGFRYSLKHFNYDICQSCFFSG RVAKGHKMHYPMVVEYCTPTTSGEDVRDFAKVLNKNKFRTKRYFAKHPRMGYLPVQTVLE GDNMET"
BASE COUNT 343 a 332 c 296 g 327 t
ORIGIN
Query Match 40.5%; Score 526.4; DB 11; Length 1298;
Best Local Similarity 93.8%; Pred. No. 8e-143;
Matches 548; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
718 CCACGAGACTCAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTT

Db 140 CCACGAGACCCAAACCACTTGTGGGACCACCCCAAAATGACAGAGCTCTACCAGTCTTT 199
Qy 778 AGCTGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCATGAAACTCGAAGACT 837
Db 200 AGCTGACCTGAATAATGTTCAGGTTCTCCGGTATAGGACTGCATGAAAGCTCAGAAGGCT 259
Qy 838 GCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTTCAGCTGCATGTGATGCCTTGGACCA 897
Db 260 CCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTTCAGCTGCATGTGATGCCTTGGACCA 319
Qy 898 GCACAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGAGATTAATAATTGTTGAC 957
Db 320 GCACAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGAGATAATTAACCTGTTGAC 379
Qy 958 CACTATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGT 1017
Db 380 TACAATTATGATCGTCTGGAGCAAGAGCACAAATCTGGTCAATGTCCCTCTCTGTGT 439
Qy 1018 GGATATGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 1077
Db 440 GGATATGTGTCTCAACTGGCTTCTCAATGTTTATGATACGGGACGAACAGGGAGGATCCG 499
Qy 1078 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA 1137
Db 500 TGTCTGTCTTTTAAACTGGCATCATTTCTCTGTGTAAAGCACACTTGGAAGACAAGTA 559
Qy 1138 CAGATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGGCTGGG 1197
Db 560 CAGATACCTTTCAAGCAAGTGGCAAGTTCAACTGGCTTTGTGACCGCAGGCTAGGCTGGG 619
Qy 1198 CCTCTCTGCAATGATTTCTATCCAAATTCGAAGACAGTTGGTGAAGTTGCATCCTTTGG 1257
Db 620 TCTTCTCTGCAATGATTTCTATTCAAATCCCAAGACAGTTGGTGAAGTTGCTTCCTTTGG 679
Qy 1258 GGGCAGTAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
Db 680 GGGCAGTAACATTGAGCCGAGTGTGAGGAGCTGCTTCCAATTTG 723
RESULT 8
AK013510 2135 bp mRNA linear HTC 05-DEC-2002
LOCUS Mus musculus adult male hippocampus cDNA, RIKEN full-length
DEFINITION enriched library, clone:290010C03 product:dystrophin, muscular dystrophy, full insert sequence.
ACCESSION AK013510
VERSION AK013510.1 GI:12850899
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,

SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
2
REFERENCE Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
AUTHORS Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
TITLE Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL 20499374
MEDLINE 11042159
PUBMED
3
REFERENCE Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
AUTHORS RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
TITLE Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL 20530913
MEDLINE 11076861
PUBMED
4
REFERENCE Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G., Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S. and Hayashizaki,Y.
AUTHORS Functional annotation of a full-length mouse cDNA collection
TITLE Nature 409 (6821), 685-690 (2001)
JOURNAL 21085660
MEDLINE 11217851
PUBMED
5
REFERENCE The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
AUTHORS Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
TITLE Nature 420, 563-573 (2002)
JOURNAL 6 (bases 1 to 4437)
REFERENCE Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,

Muramatsu,M. and Hayashizaki,Y.
Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/
FEATURES Location/Qualifiers
source 1..4437
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:9930028B14"
/db_xref="taxon:10090"
/clone="9930028B14"
/sex="female"
/tissue_type="vagina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
misc_feature 29..1968
/note="dystrophin, muscular dystrophy (MGD|MGI:94909, GB|NM_007868, evidence: BLASTN, 100%, match=999)
putative 4417..4422
/note="putative"
polyA_site 4437
/note="putative"
BASE COUNT 1300 a 932 c 885 g 1320 t
ORIGIN
Query Match 40.5%; Score 526.4; DB 11; Length 4437;
Best Local Similarity 93.8%; Pred. No. 1.8e-142;
Matches 548; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 718 CCACGAGACTCAAAACAACTTGCTGGGACCATCCCAAATGACAGAGCTCTACCAGTCTTT 777
Db 49 CCACGAGACCCAAACCACTTGTGGGACCAACCCCAAAATGACAGAGCTCTACCAGTCTTT 108
QY 778 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAGACT 837
Db 109 AGCTGACCTGAATAATGTCAGGTTCTCCGCGTATAGGACTGCCATGAAGCTCAGAAGCT 168
QY 838 GCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGTCATGTGATGCCTTGGACCA 897
Db 169 CCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGTCATGTGATGCCTTGGACCA 228
QY 898 GCACAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTTGAC 957
Db 229 GCACAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATAATTAACCTGTTTGAC 288
QY 958 CACTATTATGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTGCGT 1017
Db 289 TACAATTATGATCGTCTGGAGCAAGAGCACAACAATCTGCTCAATGTCCCTCTCTGTGT 348
QY 1018 GGATATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 1077
Db 349 GGATATGTGCTCAACTGGCTTCTCAATGTTTATGATACGGGACGAACAGGGAGGATCCG 408
QY 1078 TGTCCCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA 1137
Db 409 TGTCCCTGCTTTTAAACTGGCATCATTTCTCTGTGTAAAGCACACTTGGAAGACAAGTA 468
QY 1138 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGGCTGGG 1197

Db 469 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACTGGCTTTTGTGACCAGCGTAGGCTGGG 528

QY 1198 CCTCCTTCTGCATGATTCTATCCAAATTCCAAGACAGTGGGTGAAGTTGCATCCTTTGG 1257

Db 529 TCCTCTCTCTGCATGATTCTATTCAAATCCCAAGACAGTGGGTGAAGTTGCTTCTCTTGG 588

QY 1258 GGGCAGTAACATTGAGCCAGTGTCCGGAGCTGCTTCCAATTG 1301

Db 589 GGGCAGTAACATTGAGCCAGTGTCTCAGGAGCTGCTTCCAATTG 632

RESULT 10

AK075809

LOCUS

DEFINITION

AK075809 1384 bp mRNA linear HTC 07-DEC-2002

Mus musculus 10 day old male pancreas cDNA, RIKEN full-length enriched library, clone:1810074E05 product:dystrophin, muscular dystrophy, full insert sequence.

AK075809

AK075809.1 GI:26344655

HTC; CAP trapper.

Mus musculus (house mouse)

SOURCE

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

11042159

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

Nature 409 (6821), 685-690 (2001)

21085660

11217851

5

The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 1384)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoaka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, K., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL:http://genome.gsc.riken.go.jp/

URL:http://fantom.gsc.riken.go.jp/.

location/Qualifiers

1. 1384

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM DB:1810074E05"

/db_xref="MGI:1908112"

/db_xref="taxon:10090"

/clone="1810074E05"

/sex="male"

/tissue type="pancreas"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="10 day old"

120. 1142

/note="unnamed protein product; dystrophin, muscular dystrophy (MGD|MGI:94909, GB|NM_007868, evidence: BLASTN, 100% match=999)

putative"

/codon_start=1

/protein_id="BAC35977.1"

/db_xref="GI:26344656"

/db_xref="MGI:94909"

/translation="MREHLKGHETQTTCWDHPKMTLYQSLADLNNVRFSAYRTAMKL RRLQKALCDLDSLSAACDALDQHNKQNDQPMDLQINCLTTIYDRLEQEHNNLVN VPLQVDMCLNWLNNVDTGRTGIRVLSFKTGIISLCKAHLEDKYRLFQKQVASTGF CDQRRLGLLHDSIQIPQLGEVASFGGSNIEPSVRSFCFQANNKPEIEAALFLDWMR LE PQSMVWLPVLRHVAEAETAKHQAACNICKECPFIIGFPRYSLKHFNYDQCSCFFSG RVAXGHKMHYPMVEYCTPTTSGEDVRDFAKVLKNKFKRYFAKHPRMGYLPVQTVLE GDNMET"

BASE COUNT 376 a 354 c 303 g 351 t

ORIGIN

Query Match 40.3%; Score 524.8; DB 11; Length 1384;

Best Local Similarity 93.7%; Pred. No. 2.5e-142;

Matches 547; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```
QY 718 CCACGAGACTCAAAACAACACTTGCTGGAGCCATCCCAAAATGACAGAGCTCTACCAAGTCTTT 777
Db |||||||
QY 140 CCACGAGACCCAAACCACTTGTTGGAGCACCCCAAAATGACAGAGCTCTACCAAGTCTTT 199
Db |||||||
QY 778 AGCTGACCTGAATAATGTCAGATTCAGCTTATAGGACTGCCATGAACTCCGAGACT 837
Db |||||||
QY 200 AGCTGACCTGAATAATGTCAGGTTCTCCGGGTATAGGACTGCCATGAAGCTCAGAGGCT 259
Db |||||||
QY 838 GCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTGGACCA 897
Db |||||||
QY 260 CCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCCGGCTGCATGTGATGCCCTGGACCA 319
Db |||||||
QY 898 GCACAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTCGCAGATTATTAATTGTTGAC 957
Db |||||||
QY 320 GCACAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTCGCAGATAATTAACCTGTTGAC 379
Db |||||||
QY 958 CACTATTATGACCGCCTGGAGCAAGAGCACAAACAATTTGGTCAACGTCCTCTCTGCGT 1017
Db |||||||
QY 380 TACAATTATGATCGTCTGGAGCAAGAGCACAAACAATCTGGTCAATGTCCCTCTCTGTGT 439
Db |||||||
QY 1018 GGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGACGAACAGGGAGGATCCG 1077
Db |||||||
QY 440 GGATATGTGCTCAACTGGCTTCTCAATGTTTATGATACGGACGAACAGGGAGGATCCG 499
Db |||||||
QY 1078 TGTCCTGTCTTTAAACCTGGCATCATTTCCCTGTGTGTAAGCACATTTGGAAGACAAGTA 1137
Db |||||||
QY 500 TGTCCTGTCTTTAAACCTGGCATCATTTCTGTGTGTAAGCACACTTTGGAAGACAAGTA 559
Db |||||||
QY 1138 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGG 1197
Db |||||||
QY 560 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACTGGCTTTTGTGACCAGCGTAGGCTGGG 619
Db |||||||
QY 1198 CCTCCTTCTGCATGATTCTATCCAAATTCGAAGACAGATTGGGTGAAGTTGCATCCTTTGG 1257
Db |||||||
QY 620 TCTTCTTCTGCATGATTCTATTCAAATCCCAAGACAGATTGGGTGAAGTTGCTTCCCTTGG 679
Db |||||||
QY 1258 GGGCAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTG 1301
Db |||||||
QY 680 GGGCAGTAACATTGAGCCGAGTGTGAGGAGCTGCTTCCAAATTG 723
Db |||||||
```

```
RESULT 11
CB960722
LOCUS
DEFINITION
  AGENCOURT_13761995 NIH_MGC_147 Homo sapiens cDNA clone
  IMAGE:30344479 5', mRNA sequence.
CB960722
CB960722.1 GI:30216839
EST.
SOURCE
  Homo sapiens (human)
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  NIH-MGC http://mgi.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished
  Contact: Robert Strausberg, Ph.D.
  Email: cgabs-r@mail.nih.gov
  Tissue Procurement: Dr. Stefan Hansson
  cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
  and advice from Piero Carninci (RIKEN)
  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Agencourt Bioscience Corporation
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  Plate: NDAM380 row: n column: 08
  High quality sequence stop: 620.
  Location/Qualifiers
    1..797
    /organism="Homo sapiens"
```

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30344479"
/tissue_type="Human Placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_147"
/note="Organ: placenta; Vector: pBluescriptR; Site: 1:
all-XhoI; Site 2: BamH; Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTTN-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIMH/NHGRI, National Institutes of Health). Note: This is
a NIH MGC library."
```

```
BASE COUNT      199 a      208 c      199 g      191 t
ORIGIN

Query Match      39.3%; Score 511; DB 14; Length 797;
Best Local Similarity 99.8%; Pred. No. 1.9e-138;
Matches 522; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 718 CCACGAGACTCAAAACAACCTTGCTGGAGCCATCCCAAAATGACAGAGCTCTACCAAGTCTTT 777
Db |||||||
QY 168 CCACGAGACTCAAAACAACCTTGCTGGAGCCATCCCAAAATGACAGAGCTCTACCAAGTCTTT 227
Db |||||||
QY 778 AGCTGACCTGAATAATGTCAGATTCAGCTTATAGGACTGCCATGAACTCCGAGACT 837
Db |||||||
QY 228 AGCTGACCTGAATAATGTCAGATTCAGCTTATAGGACTGCCATGAACTCCGAGACT 287
Db |||||||
QY 838 GCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTGGACCA 897
Db |||||||
QY 288 GCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTGGACCA 347
Db |||||||
QY 898 GCACAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTGAC 957
Db |||||||
QY 348 GCACAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTGAC 407
Db |||||||
QY 958 CACTATTATGACCGCCTGGAGCAAGAGCACAAACAATTTGGTCAACGTCCTCTCTGCGT 1017
Db |||||||
QY 408 CACTATTATGACCGCCTGGAGCAAGAGCACAAACAATTTGGTCAACGTCCTCTCTGCGT 467
Db |||||||
QY 1018 GGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGACGAACAGGGAGGATCCG 1077
Db |||||||
QY 468 GGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGACGAACAGGGAGGATCCG 527
Db |||||||
QY 1078 TGTCCTGTCTTTTAAACCTGGCATCATTTCCCTGTGTGTAAGCACATTTTGGAGACAAGTA 1137
Db |||||||
QY 528 TGTCCTGTCTTTTAAACCTGGCATCATTTCCCTGTGTGTAAGCACATTTTGGAGACAAGTA 587
Db |||||||
QY 1138 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGG 1197
Db |||||||
QY 588 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGG 647
Db |||||||
QY 1198 -CCTCCTTCTGCATGATTCTATCCAAATTCGAAGACAGATTGGG 1239
Db |||||||
QY 648 CCTCCTTCTGCATGATTCTATCCAAATTCGAAGACAGATTGGG 690
Db |||||||
```

```
RESULT 12
BB629984
LOCUS
DEFINITION
  BB629984 RIKEN full-length enriched, adult female vagina Mus
  musculus cDNA clone 9930028B14 5', mRNA sequence.
BB629984
BB629984.1 GI:15399733
EST.
SOURCE
  Mus musculus (house mouse)
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
  1 (bases 1 to 652)
  Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
```


primer CGACCTGCAGCTCGACACA."

202 a 232 c 208 g 209 t 3 others

Query Match 36.9%; Score 479.6; DB 9; Length 854;
Best Local Similarity 90.9%; Pred.No. 3.4e-129;
Matches 531; Conservative 0; Mismatches 51; Indels 2; Gaps 2;
QY 718 CCACGAGACTCAAACTTGTCTGGGACCATCCCAAATGACAGAGCTCTACCAGTCTTT 777
Db 135 CCACGAGACCCAAACCACTTGTGGGACCAACCCCAAATGACAGAGCTCTACCAGTCTTT 194
QY 778 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACT 837
Db 195 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGCT 254
QY 838 GCAGAAAGGCCCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCTTGGACCA 897
Db 255 CCAGAAAGGCCCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCTTGGACCA 314
QY 898 GCACAACTCAAGCAAAATGACAGCCCATGGATATCCTGCGAGATTATTAATTGTTGAC 957
Db 315 GCACAACTCAAGCAAAATGACAGCCCATGGATATCCTGCGAGATTATTAATTGTTGAC 374
QY 958 CACTATTTATGACCGCCTGGAGCAAGAGACAAATTTGGTCAAGCTCCCTCTCTGGCT 1017
Db 375 TACAATTTATGATCGTCTGGAGCAAGAGACAAATCTGGTCAATGTCCCTCTCTGTGT 434
QY 1018 GGATATGHTCTGAAGTGGCTGTGAATGTTTATGATACGGGACGAACAGGAGGATCC- 1076
Db 435 GGATATGHTCTGAAGTGGCTTCTCAATGTTTATGATACGGGACGAACAGGAGATCCG 494
QY 1077 GTGTCTCTGCTTTTAAACTGGCATCAATTCCTGTGTAAAGCAATTTGGAAGACAAGT 1136
Db 495 GGGTCTCTGCTTTTAAACTGGCATCAATTCCTGTGTAAAGCAATTTGGAAGACAAGT 554
QY 1137 ACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACGCGCAGGCTGG 1196
Db 555 ACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACGCGTANGCTGC 614
QY 1197 GCCTCCTTCTGCATGATCTATCCAAATTCGAAGACAGTTGGTGAAGTTGCATCCTTTG 1256
Db 615 GTCTTCTTCTGCATGATCTATTACATCCCAAGACAGGTTGGTGAAGTGCTTCCTTT- 673
QY 1257 GGGGCAGTAACATTGACCAAGTGTCGGGAGTGCTTCCAATTT 1300
Db 674 GGGGCAGTAACATTGACCGAGTGTCAGGAGTGCTTCCACTTT 717

RESULT 15
CB527785
LOCUS CB527785 717 bp mRNA linear EST 28-MAR-2003
DEFINITION UI-M-FY0-cfk-c-14-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
IMAGE: 6850599 5', mRNA sequence.
ACCESSION CB527785
VERSION CB527785.1 GI:29361258
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 717)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousef1.html>

This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.

Location/Qualifiers

source

1. .717
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6850599"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FY0"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is ACCGAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 180 a 167 c 171 g 198 t 1 others
ORIGIN

Query Match 36.2%; Score 470.8; DB 14; Length 717;
Best Local Similarity 93.9%; Pred.No. 1.2e-126;
Matches 490; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 780 CTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGC 839
Db 1 CTGACCTGAATAATGTCAGATTCTCCGCGTATAGGACTGCCATGAACTCAGAAGGCTCC 60
QY 840 AGAAGGCCCTTGTGCTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCTTGGACCA 899
Db 61 AGAAGGCCCTTGTGCTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCTTGGACCA 120
QY 900 ACAACCTCAAGCAAAATGACCGCCATGGATATCTGCGAGATTATTAATTGTTGACCA 959
Db 121 ACAACCTCAAGCAAAATGACCGCCATGGATATCTGCGAGATTATTAATTGTTGACTA 180
QY 960 CTATTTATGACCGCCTGGAGCAAGACACAACAATTTGGTCAAGCTCCCTCTCTGCGTGG 1019
Db 181 CAATTTATGATCGTCTGGAGCAAGAGCACAAATCTGGTCAATGTCCCTCTCTGCTGG 240
QY 1020 ATATGTGCTCGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTG 1079
Db 241 ATATGTGCTCGAACTGGCTTCTCAATGTTTATGATACGGGACGAACAGGAGGATCCGTG 300
QY 1080 TCCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACA 1139
Db 301 TCCTGTCTTTTAAACTGGCATCATTTCTGTGTAAAGCACATTTGGAAGACAAGTACA 360
QY 1140 GATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACGAGCGAGGCTGGGCC 1199
Db 361 GATACCTTTTCAAGCAAGTGGCAAGTTCAACTGGCTTTTGTGACGAGCGTAGGCTGGGTC 420
QY 1200 TCCTTCTGCATGATTTCTATCCAAATTCGAAGACAGTTGGTGAAGTTGTCATCCTTTGGG 1259
Db 421 TTCTTCTGCATGATTTCTATTCAAATCCCAAGACAGTTGGTGAAGTTGCTTCTTTGGG 480
QY 1260 GCAGTAACATTGACCCAAAGTGTCCGGAGCTGCTTCCAATTG 1301
Db 481 GCAGTAACATTGACCCGAGTGTGAGGAGTGTCTTCCAATTG 522

Search completed: February 2, 2004, 06:03:15
Job time : 2442.84 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: February 1, 2004, 10:14:55 ; Search time 7484.44 Seconds
(without alignments)
11483.979 Million cell updates/sec
Title: US-09-845-416-12_COPY_900_3000
Perfect score: 2101
Sequence: 1 gagctatgcctacacacagg.....ttctgcatgattctatccaa 2101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID | Description |
|------------|--------|-------|--------|----|-----------|--------------------|
| 1 | 1760.4 | 83.8 | 5339 | 6 | AX538620 | AX538620 Sequence |
| 2 | 1629 | 77.5 | 5462 | 6 | AX538621 | AX538621 Sequence |
| 3 | 1611.6 | 76.7 | 5417 | 6 | AX538619 | AX538619 Sequence |
| 4 | 1328.8 | 63.2 | 5952 | 6 | AR304538 | AR304538 Sequence |
| 5 | 1328.8 | 63.2 | 5952 | 6 | AX114289 | AX114289 Sequence |
| 6 | 1328.8 | 63.2 | 8689 | 6 | AX538622 | AX538622 Sequence |
| 7 | 1328.8 | 63.2 | 11443 | 6 | AX538624 | AX538624 Sequence |
| 8 | 1328.8 | 63.2 | 12057 | 6 | AX538627 | AX538627 Sequence |
| 9 | 1328.8 | 63.2 | 12446 | 9 | HSDMDR | X14298 Human mRNA |
| 10 | 1328.8 | 63.2 | 13957 | 6 | AX409637 | AX409637 Sequence |
| 11 | 1328.8 | 63.2 | 13957 | 6 | AX538581 | AX538581 Sequence |
| 12 | 1328.8 | 63.2 | 13957 | 9 | HUMDYS | M18533 Homo sapien |
| 13 | 1317.8 | 62.7 | 13977 | 6 | AR220819 | AR220819 Sequence |
| 14 | 1317 | 62.7 | 4402 | 6 | E30219 | E30219 Shortened d |
| 15 | 1196.6 | 57.0 | 4402 | 6 | E30220 | E30220 Shortened d |
| 16 | 1185.4 | 56.4 | 13887 | 4 | AF070485 | AF070485 Canis fam |
| 17 | 1149.6 | 54.7 | 13815 | 6 | AX306153 | AX306153 Sequence |
| 18 | 1149.6 | 54.7 | 13815 | 6 | AX538582 | AX538582 Sequence |
| 19 | 1149.6 | 54.7 | 13815 | 10 | MUSDYSA | M68859 Mouse dystr |
| 20 | 1149.6 | 54.7 | 19307 | 6 | AR093392 | AR093392 Sequence |
| 21 | 1149.6 | 54.7 | 19307 | 6 | AR142592 | AR142592 Sequence |
| 22 | 922.2 | 43.9 | 13575 | 5 | GGDYS | X13369 Chicken mRN |
| 23 | 718.8 | 34.2 | 4075 | 6 | E30221 | E30221 Shortened d |
| 24 | 715.6 | 34.1 | 2654 | 5 | FSCDYSTRO | M37645 Torpedo cal |
| 25 | 667.2 | 31.8 | 3747 | 6 | E30218 | E30218 Shortened d |
| 26 | 629.8 | 30.0 | 3275 | 10 | MUSDYS | M18025 Mouse dystr |
| 27 | 613.6 | 29.2 | 3163 | 6 | E30223 | E30223 Shortened d |
| 28 | 598.2 | 28.5 | 3521 | 5 | AF339031 | AF339031 Danio rer |
| 29 | 555.6 | 26.4 | 10705 | 10 | RNAJ2967 | AJ002967 Rattus no |
| 30 | 554.2 | 26.4 | 11096 | 6 | AX538584 | AX538584 Sequence |
| 31 | 554.2 | 26.4 | 11096 | 10 | MMY12229 | Y12229 M.musculus |
| 32 | 554 | 26.4 | 3161 | 10 | MMGUTRPH | X83506 M.musculus |
| 33 | 547.4 | 26.1 | 6045 | 6 | A63605 | A63605 Sequence 7 |
| 34 | 547.4 | 26.1 | 6045 | 6 | AR281528 | AR281528 Sequence |
| 35 | 547.4 | 26.1 | 6059 | 6 | AX107972 | AX107972 Sequence |
| 36 | 547.4 | 26.1 | 10302 | 6 | AX538583 | AX538583 Sequence |
| 37 | 547.4 | 26.1 | 10302 | 9 | HSMUPS | X69086 H.sapiens m |
| 38 | 547.4 | 26.1 | 10320 | 6 | A63607 | A63607 Sequence 9 |
| 39 | 547.4 | 26.1 | 10320 | 6 | AR281529 | AR281529 Sequence |
| 40 | 528.4 | 25.1 | 3073 | 10 | AF195788 | AF195788 Rattus no |
| 41 | 528.4 | 25.1 | 3140 | 10 | AF195787 | AF195787 Rattus no |
| 42 | 527.4 | 25.1 | 3499 | 9 | HSU43519 | U43519 Human dystr |
| 43 | 527.4 | 25.1 | 5106 | 6 | AX552248 | AX552248 Sequence |
| 44 | 505 | 24.0 | 2110 | 9 | HUMDMDX | M92650 Humar Duche |
| 45 | 505 | 24.0 | 4658 | 9 | BC028720 | BC028720 Homo sapi |

ALIGNMENTS

RESULT 1
AX538620
LOCUS AX538620 5339 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 40 from Patent WO0229056.
ACCESSION AX538620
VERSION AX538620.1 GI:25271166
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Chamberlain,J.S. and Harper,S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 40 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

| FEATURES | | Location/Qualifiers | |
|--|-----------------------------|---|------|
| source | 1..5339 | /organism="synthetic construct" | |
| | | /mol_type="genomic DNA" | |
| | | /db_xref="taxon:32630" | |
| | | /note="Synthetic" | |
| BASE COUNT | 1638 a 1191 c 1187 g 1323 t | | |
| ORIGIN | | | |
| Query Match 83.8%; Score 1760.4; DB 6; Length 5339; | | | |
| Best Local Similarity 90.3%; Pred. No. 0; | | | |
| Matches 1910; Conservative 0; Mismatches 191; Indels 15; Gaps 2; | | | |
| QY | 1 | GAGCTATGCCTACACACAGGCTGCTTATGTTCACCACCTCTGACCCCTACACGGAGCCCAT | 60 |
| Db | 1099 | GAGCTATGCCTACACACAGGCTGCTTATGTTCACCACCTCTGACCCCTACACGGAGCCCAT | 1158 |
| QY | 61 | TCCTTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTTCATTGATGAGAG | 120 |
| Db | 1159 | TCCTTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTTCATTGATGAGAG | 1218 |
| QY | 121 | TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC | 180 |
| Db | 1219 | TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC | 1278 |
| QY | 181 | TGCTGAGGACACATTTCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA | 240 |
| Db | 1279 | TGCTGAGGACACATTTCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA | 1338 |
| QY | 241 | CCAGTTTCATACTGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCCGGTTGG | 300 |
| Db | 1339 | CCAGTTTCATACTGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCCGGTTGG | 1398 |
| QY | 301 | TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA | 360 |
| Db | 1399 | TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA | 1458 |
| QY | 361 | AACTGAAGTACAAGACGAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC | 420 |
| Db | 1459 | AACTGAAGTACAAGACGAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC | 1518 |
| QY | 421 | TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT | 480 |
| Db | 1519 | TAGCATGGAAAAACAAGCAATTTACATCATAGATTTACTGCAACAGTTCCTCCCTGGACCT | 1578 |
| QY | 481 | GAAAGAGTTGAATGACTGGCTAAACAAAAACAAGAAAGAAACAAGGAAATGGAGGAAGA | 540 |
| Db | 1579 | GGAAAGTTTCTTGCCTGGCTTACAGAGCTGAAACAACTGCCAATGTCTCTACAGGATGC | 1638 |
| QY | 541 | GCCTCTTGGACCTGATCTTTGAAGACCTAAACCGCAAGTACAACAATAAGGTGCTTCA | 600 |
| Db | 1639 | TACCCGTAAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATG | 1698 |
| QY | 601 | AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTAGT | 660 |
| Db | 1699 | GCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAAA | 1758 |
| QY | 661 | TGATGA-----ATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAAACAAC | 708 |
| Db | 1759 | CAGCAAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTTACAAAGACG | 1818 |
| QY | 709 | TAAGG---TATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGGGT | 765 |
| Db | 1819 | TTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGGTC | 1878 |
| QY | 766 | TCCTTTTACAAGACAGTTCTGACAGTGAAGCGCTGACACCTTTTCTCTGCAGGAACCTTCT | 825 |
| Db | 1879 | CCATTTGGAAGCCAGTTCTGACAGTGAAGCGTCTGACCTTTTCTCTGCAGGAACCTTCT | 1938 |
| QY | 826 | GGTGTGGTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCGGACTT | 885 |
| Db | 1939 | GGTGTGGTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCGGACTT | 1998 |
| QY | 886 | TCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGCGCTTCAAGAGGGAAATTGAAAACTAA | 945 |
| Db | 1999 | TCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGCGCTTCAAGAGGGAAATTGAAAACTAA | 2058 |
| QY | 946 | AGAACCTGTAAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTT | 1005 |
| Db | 2059 | AGAACCTGTAAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTT | 2118 |
| QY | 1006 | GGAAGGACTAGAGAAAACCTTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCA | 1065 |
| Db | 2119 | GGAAGGACTAGAGAAAACCTTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCA | 2178 |
| QY | 1066 | GAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATT | 1125 |
| Db | 2179 | GAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATT | 2238 |
| QY | 1126 | GAACCTGCACTCCGCTGACTGGCAGAGAAAAAATAGATGAGACCCCTTGAAGAATCCAGGA | 1185 |
| Db | 2239 | GAACCTGCACTCCGCTGACTGGCAGAGAAAAAATAGATGAGACCCCTTGAAGAATCCAGGA | 2298 |
| QY | 1186 | ACTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGGGCAAGCTGAGGTGATCAAGG | 1245 |
| Db | 2299 | ACTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGGGCAAGCTGAGGTGATCAAGG | 2358 |
| QY | 1246 | ATCCTGGCAGCCCGTGGCGATCTCCTCATTTGATGACTCTCTCAAGATCACCTCGAGAAAGT | 1305 |
| Db | 2359 | ATCCTGGCAGCCCGTGGCGATCTCCTCATTTGATGACTCTCTCAAGATCACCTCGAGAAAGT | 2418 |
| QY | 1306 | CAAGGCACCTTCGAGGAGAAATTGCGCTCTGAAAGAGAACGAGAGCCACGTCATGACCT | 1365 |
| Db | 2419 | CAAGGCACCTTCGAGGAGAAATTGCGCTCTGAAAGAGAACGAGAGCCACGTCATGACCT | 2478 |
| QY | 1366 | TGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTGGA | 1425 |
| Db | 2479 | TGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTGGA | 2538 |
| QY | 1426 | AGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCT | 1485 |
| Db | 2539 | AGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCT | 2598 |
| QY | 1486 | GCATGAAGCCCAACAGGGAATTTGGTCCAGCATCTCAGCACCTTTTCCACGCTCTGTCCA | 1545 |
| Db | 2599 | GCATGAAGCCCAACAGGGAATTTGGTCCAGCATCTCAGCACCTTTTCCACGCTCTGTCCA | 2658 |
| QY | 1546 | GGGTCCCTGGGAGAGGCCATCTCGCCAAAAAAGTGCCCTACTATATCAACCACGAGAC | 1605 |
| Db | 2659 | GGGTCCCTGGGAGAGGCCATCTCGCCAAAAAAGTGCCCTACTATATCAACCACGAGAC | 2718 |
| QY | 1606 | TCAAACAACCTTGTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCT | 1665 |
| Db | 2719 | TCAAACAACCTTGTGTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCT | 2778 |
| QY | 1666 | GAATAATGTGAGATTTCTCAGCTTATAGGACTGCCATGAAAACCTCCGAAGACTGCAGAAGGC | 1725 |
| Db | 2779 | GAATAATGTGAGATTTCTCAGCTTATAGGACTGCCATGAAAACCTCCGAAGACTGCAGAAGGC | 2838 |
| QY | 1726 | CCTTTGCTTGGATCTCTGAGCCTGTGAGCTGCAATGATGCCTTGGACCAACACCT | 1785 |
| Db | 2839 | CCTTTGCTTGGATCTCTGAGCCTGTGAGCTGCAATGATGCCTTGGACCAACACCT | 2898 |
| QY | 1786 | CAAGCAAAATGACAGCCCATGGATATCCTGAGATTTAATTTGTTGACCACTATTTA | 1845 |
| Db | 2899 | CAAGCAAAATGACAGCCCATGGATATCCTGAGATTTAATTTGTTGACCACTATTTA | 2958 |
| QY | 1846 | TGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTCGGTGGATATGTG | 1905 |
| Db | 2959 | TGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTCGGTGGATATGTG | 3018 |
| QY | 1906 | TCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGCTCTGTC | 1965 |
| Db | 3019 | TCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGCTCTGTC | 3078 |
| QY | 1966 | TTTTAAAACTGGCATCATTTCCCTGTGTGTAAGACACATTTGGAAGACAAGTACAGATACCT | 2025 |

```
Db 3079 TTTTAAACTGCGCATCATTTCCCTGTGTAAAGCACATTTGGAAAGACAAGTACAGATACCT 3138
QY 2026 TTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCTCCTTCT 2085
Db 3139 TTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCTCCTTCT 3198
QY 2086 GCATGATTCATCCAA 2101
Db 3199 GCATGATTCATCCAA 3214

RESULT 2
AX538621
LOCUS AX538621 5462 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 41 from Patent WO0229056.
ACCESSION AX538621
VERSION AX538621.1 GI:25271168
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE
1 Chamberlain,J.S. and Harper,S.Q.
Mini-dystrophin nucleic acid and peptide sequences
Patent: WO 0229056-A 41 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
1..5462
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"
BASE COUNT 1668 a 1225 c 1212 g 1357 t
ORIGIN
Query Match 77.5%; Score 1629; DB 6; Length 5462;
Best Local Similarity 85.4%; Pred.No. 0;
Matches 1911; Conservative 0; Mismatches 190; Indels 138; Gaps 3;
QY 1 GAGCTATGCCATACACACAGGCTGCTTATGTCCACCCTCTGACCCTACACGGAGCCCAT 60
Db 1099 GAGCTATGCCATACACACAGGCTGCTTATGTCCACCCTCTGACCCTACACGGAGCCCAT 1158
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCAATGATGAGAG 120
Db 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCAATGATGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATATCGTGGCTTCTTTC 180
Db 1219 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATATCGTGGCTTCTTTC 1278
QY 181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
QY 241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
Db 1339 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1398
QY 301 TAATATTCTACAAATTGGGAAGTAAGCTGATGGAAACAGGAAATATATCAGAAGATGAAGA 360
Db 1399 TAATATTCTACAAATTGGGAAGTAAGCTGATGGAAACAGGAAATATATCAGAAGATGAAGA 1458
QY 361 AACTGAAGTACAGACAGATGAATCTCCTAAATTCAAGATGGGAATGCCCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAGACAGATGAATCTCCTAAATTCAAGATGGGAATGCCCTCAGGGTAGC 1518
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTT----- 454
Db 1519 TAGCATGGAAAAACAAAGCAATTTACATGCTCCTGGACTGACCACCTATTGGAGCCTCTCC 1578
QY 455 ----- 454
```

```
Db 1579 TACTCAGACTGTTACTCTGTGTGACACAACCTGTGGTTACTTAAGGAAAACGCCATCTCCAA 1638
QY 455 -----TTAATGGATCTCCAGAATCAGAA 477
Db 1639 ACTAGAAATGCCATCTTCCTTGATGTGGAGCATAGATTACTGCAACAGTTCCCCCTGGA 1698
QY 478 ACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAAATGGAGGA 537
Db 1699 CCTGGAAAAGTTTCTTGCCCTGGCTTACAGAAAGCTGAAACAACCTGCCAATGTCTCTACAGGA 1758
QY 538 AGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAACATAAAGGTGCT 597
Db 1759 TGCTACCCGTAAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACA 1818
QY 598 TCAAGAAGATCTAGAACAAGAACAAGTCAAGGTCAATTCTCTCACTCACATGGTGGTGGT 657
Db 1819 ATGGCAAGACCTCCCAAGGTGAATTGAAGCTCACACAGATGTTATCACAACCTGGATGA 1878
QY 658 AGTTGATGA-----ATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGACA 705
Db 1879 AAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGAGTCTCTTTACAAAG 1938
QY 706 ACTTAAG---TATTGGGAGATCGATGGGCAACATCTGTATGAGACAGAAAGACCCGCTG 762
Db 1939 ACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAG 1998
QY 763 GGTTCCTTTACAAGACAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACT 822
Db 1999 GTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACT 2058
QY 823 TCTGTGTGCTACAGCTGAAAGATGATGAATTAAAGCCGGCAGGCACCTATTGGAGGCGGA 882
Db 2059 TCTGTGTGCTACAGCTGAAAGATGATGAATTAAAGCCGGCAGGCACCTATTGGAGGCGGA 2118
QY 883 CTTTCCAGCAGTTCAGAAAGCAGAACCATGTACATAGGGCCCTTCAAGAGGGAATTTGAAAAC 942
Db 2119 CTTTCCAGCAGTTCAGAAAGCAGAACCATGTACATAGGGCCCTTCAAGAGGGAATTTGAAAAC 2178
QY 943 TAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCC 1002
Db 2179 TAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCC 2238
QY 1003 TTTTGAAGGACTAGAGAAACTCTACAGAGGAGCCAGAGAGTGCCTCTCTGAGGAGAGAGC 1062
Db 2239 TTTTGAAGGACTAGAGAAACTCTACAGAGGAGCCAGAGAGTGCCTCTCTGAGGAGAGAGC 2298
QY 1063 CCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGAAAA 1122
Db 2299 CCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGAAAA 2358
QY 1123 ATTGAACCTGCACCTCCGCTGACTGSCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCA 1182
Db 2359 ATTGAACCTGCACCTCCGCTGACTGSCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCA 2418
QY 1183 GGAACCTTCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAA 1242
Db 2419 GGAACCTTCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAA 2478
QY 1243 GGGATCCTGGCAGCCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAA 1302
Db 2479 GGGATCCTGGCAGCCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAA 2538
QY 1303 AGTCAAGGCACTTCGAGGAGAAATTCGCCCTCTGAAAAGAGAACGTCGAGCCACGTCAATGA 1362
Db 2539 AGTCAAGGCACTTCGAGGAGAAATTCGCCCTCTGAAAAGAGAACGTCGAGCCACGTCAATGA 2598
QY 1363 CCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTACCGTATAACCTCAGCACTCT 1422
Db 2599 CCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTACCGTATAACCTCAGCACTCT 2658
QY 1423 GGAAGACCTTGAACACACAGATGGAAGCTTCTGCAGGTGGCGTCGAGGACCGAGTCAGGCA 1482
```


| | | | |
|--|------|--|------|
| Db | 2659 | GGAAGACCTGAACACCAGATGGAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCA | 2718 |
| QY | 1483 | GCTGCATGAAGCCACACAGGACITTTGGTCCAGCATCTCAGCACITTTCTTTCCACGTCGT | 1542 |
| Db | 2719 | GCTGCATGAAGCCACACAGGACITTTGGTCCAGCATCTCAGCACITTTCTTTCCACGTCGT | 2778 |
| QY | 1543 | CCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCACGA | 1602 |
| Db | 2779 | CCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCACGA | 2838 |
| QY | 1603 | GACTCAAAACAACCTTGCTGGGACCAATCCCAAAATGACAGAGCTCTACCACTTTTAGCTGA | 1662 |
| Db | 2839 | GACTCAAAACAACCTTGCTGGGACCAATCCCAAAATGACAGAGCTCTACCACTTTTAGCTGA | 2898 |
| QY | 1663 | CCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAA | 1722 |
| Db | 2899 | CCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAA | 2958 |
| QY | 1723 | GGCCCTTTGCTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGAGACCAGCAA | 1782 |
| Db | 2959 | GGCCCTTTGCTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGAGACCAGCAA | 3018 |
| QY | 1783 | CCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTTTGACCACTAT | 1842 |
| Db | 3019 | CCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTTTGACCACTAT | 3078 |
| QY | 1843 | TTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGATAT | 1902 |
| Db | 3079 | TTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGATAT | 3138 |
| QY | 1903 | GTGTCTGAACCTGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCCT | 1962 |
| Db | 3139 | GTGTCTGAACCTGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCCT | 3198 |
| QY | 1963 | GTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATA | 2022 |
| Db | 3199 | GTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATA | 3258 |
| QY | 2023 | CCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGSCCTCCT | 2082 |
| Db | 3259 | CCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGSCCTCCT | 3318 |
| QY | 2083 | TCTGCATGATTTCTATCCAA | 2101 |
| Db | 3319 | TCTGCATGATTTCTATCCAA | 3337 |
| RESULT 3 | | | |
| AX538619 | | | |
| LOCUS | | | |
| Sequence 39 from Patent WO0229056. | | | |
| AX538619 | | | |
| VERSION | | | |
| AX538619.1 GI:25271163 | | | |
| KEYWORDS | | | |
| synthetic construct | | | |
| SOURCE | | | |
| synthetic construct | | | |
| ORGANISM | | | |
| artificial sequences. | | | |
| REFERENCE | | | |
| 1 | | | |
| AUTHORS | | | |
| Chamberlain, J.S. and Harper, S.O. | | | |
| TITLE | | | |
| Mini-dystrophin nucleic acid and peptide sequences | | | |
| JOURNAL | | | |
| Patent: WO 0229056-A 39 11-APR-2002; | | | |
| THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US) | | | |
| FEATURES | | | |
| Location/Qualifiers | | | |
| 1..5417 | | | |
| /organism="synthetic construct" | | | |
| /mol_type="genomic DNA" | | | |
| /db_xref="taxon:32630" | | | |
| /note="Synthetic" | | | |
| BASE COUNT | | | |
| 1700 a 1192 c 1182 g 1343 t | | | |
| ORIGIN | | | |
| Query Match | | | |
| 76.7%; Score 1611.6; DB 6; Length 5417; | | | |
| Best Local Similarity | | | |
| 85.3%; Pred. No. 0; | | | |

| | | | |
|--|------|--|------|
| Matches 1872; Conservative 0; Mismatches 229; Indels 93; Gaps 3; | | | |
| QY | 1 | GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACTCTGACCACTACACGGAGCCCAAT | 60 |
| Db | 1099 | GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACTCTGACCACTACACGGAGCCCAAT | 1158 |
| QY | 61 | TCCTTCACAGCATTGTGAAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCAATTGATGGAGAG | 120 |
| Db | 1159 | TCCTTCACAGCATTGTGAAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCAATTGATGGAGAG | 1218 |
| QY | 121 | TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC | 180 |
| Db | 1219 | TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC | 1278 |
| QY | 181 | TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCCTAATGATGTGGAAGTGGTGAAGA | 240 |
| Db | 1279 | TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCCTAATGATGTGGAAGTGGTGAAGA | 1338 |
| QY | 241 | CCAGTTTCATACCTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTGG | 300 |
| Db | 1339 | CCAGTTTCATACCTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTGG | 1398 |
| QY | 301 | TAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA | 360 |
| Db | 1399 | TAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA | 1458 |
| QY | 361 | AACTGAACTACAAGAGCAGATGAATCTCTTAAATTCACAGTGGGAATGCCTCAGGCTAGC | 420 |
| Db | 1459 | AACTGAACTACAAGAGCAGATGAATCTCTTAAATTCACAGTGGGAATGCCTCAGGCTAGC | 1518 |
| QY | 421 | TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT | 480 |
| Db | 1519 | TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT | 1578 |
| QY | 481 | GAAGAGTTGAATGACTGGCTTAACAAAAACAGAGAAGAAACAGGAAAAATGGAGGAAGA | 540 |
| Db | 1579 | GAAGAGTTGAATGACTGGCTTAACAAAAACAGAGAAGAAACAGGAAAAATGGAGGAAGA | 1638 |
| QY | 541 | GCCTCTTGGACCTGATCTTGAGACCTTAAACGCCAAAGTACAAACATTAAGTGCTTCA | 600 |
| Db | 1639 | GCCTCTTGGACCTGATCTTGAGACCTTAAACGCCAAAGTACAAACATTAAGTGCTTCA | 1698 |
| QY | 601 | AGAAGATCTAGAACAAGAACAAAGTCAGGTCATCTCTCACTCACATGGTGGTGTAGT | 660 |
| Db | 1699 | AGAAGATCTAGAACAAGAACAAAGTCAGGTCATCTCTCACTCACATGGTGGTGTAGT | 1758 |
| QY | 661 | TGATGAATCTAGTGGAGATCACGCCAACTGCTGCTTTGGAAGAAACAACTTAAGTTATGGG | 720 |
| Db | 1759 | TGATGAATCTAGTGGAGATCACGCCAACTGCTGCTTTGGAAGAAACAACTTAAGTTATGGG | 1818 |
| QY | 721 | AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGTCTCTTTTACAGACAG | 780 |
| Db | 1819 | AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGTCTCTTTTACAGACAT | 1878 |
| QY | 781 | TTCTGACCACTGGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCT | 840 |
| Db | 1879 | CCTTCTCAAATGGCAACGTCCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA | 1938 |
| QY | 841 | GAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGCGACTTTCAGCAGTTCAGAA | 900 |
| Db | 1939 | AAAAGAAGATGCAGTGAAACAAGATTCACACAACACTGGCTTTTAAAGATCAAATGAATGTT | 1998 |
| QY | 901 | GCAGAACGATCTACATAGGGCC--TTCAAGAGGGAAATTGAAAACTAAAGAACCTGTAATC | 958 |
| Db | 1999 | ATCAAGTCTTCAAAAACTGGCCCGTTTAAAGCGGATCTAGAAAAAGAAAAGCAATCCAT | 2058 |
| QY | 959 | ATGAGTACTCTTGAGACTGTACGAATATTT-----CTGACAGAGCAGCCTTTTGA | 1008 |
| Db | 2059 | GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC | 2118 |
| QY | 1009 | AGGACTAGAGAAACTCTACAGGAGCCACAGAGCTGCCTCTCTGAGGAGAGAGCCCGAGAA | 1068 |
| Db | 2119 | CCAGAAGACGGAAGCATGGCTGGATACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA | 2178 |

QY 1069 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATAC----- 1110
Db 2179 ACTTGAAGAGTACAGCACAGATTTTCACAGGCTGTCCACCACACTCAGCCATCACTAAC 2238
QY 1111 ----- 1110
Db 2239 ACAGACAACCTGAATGGAACAGTAACTACGGTGACCAACAGGGAACAGATCCTGGTAAA 2298
QY 1111 ---TGAGTGGAAAAATTGAACCTGCACTCCGCTGACTGCCAGAGAAAAATAGATGAGAC 1167
Db 2299 GCATGCTCAAGAGAACTTCCACCACCTCCCAAAAGAGAGGCAGATTACTGTGA 2358
QY 1168 CTTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCA 1227
Db 2359 TCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCA 2418
QY 1228 AGCTGAGGTGATCAAGGGATCCTGGCAGCCCCGTGGCGATCTCCTCATTGACTCTCTCCA 1287
Db 2419 AGCTGAGGTGATCAAGGGATCCTGGCAGCCCCGTGGCGATCTCCTCATTGACTCTCTCCA 2478
QY 1288 AGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGAAATTGGCGCTCTGAAAGAGAACGT 1347
Db 2479 AGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGAAATTGGCGCTCTGAAAGAGAACGT 2538
QY 1348 GAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGATTCAGCTCTCACCGTA 1407
Db 2539 GAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGATTCAGCTCTCACCGTA 2598
QY 1408 TAACCTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTGA 1467
Db 2599 TAACCTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTGA 2658
QY 1468 GGACCGAGTCAGGAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTT 1527
Db 2659 GGACCGAGTCAGGAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTT 2718
QY 1528 TCTTTCACGTCTGTCCAGGGTCCCTGGGAGAGGCCATCTCGCAAAACAAAGTGCCCTA 1587
Db 2719 TCTTTCACGTCTGTCCAGGGTCCCTGGGAGAGGCCATCTCGCAAAACAAAGTGCCCTA 2778
QY 1588 CTATATCAACCAGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTA 1647
Db 2779 CTATATCAACCAGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTA 2838
QY 1648 CCAGTCTTTAGCTGACCTGAATAATGTTCAGATTCTCAGCTTATAGACTGCCATGAAACT 1707
Db 2839 CCAGTCTTTAGCTGACCTGAATAATGTTCAGATTCTCAGCTTATAGACTGCCATGAAACT 2898
QY 1708 CCGAAGACTGCAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTTCAGCTGCATGTGATGC 1767
Db 2899 CCGAAGACTGCAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTTCAGCTGCATGTGATGC 2958
QY 1768 CTTGGACCAAGCAACCTCAAGCAAAATGACAGCCCCATGGATATCCTGCAGATTATTAA 1827
Db 2959 CTTGGACCAAGCAACCTCAAGCAAAATGACAGCCCCATGGATATCCTGCAGATTATTAA 3018
QY 1828 TTGTTTGAACCACTATTATGACCGCCTGGAGCAAGAGCAACAATTGTTGGTCAACGTCCC 1887
Db 3019 TTGTTTGAACCACTATTATGACCGCCTGGAGCAAGAGCAACAATTGTTGGTCAACGTCCC 3078
QY 1888 TCTCTGCGTGGATATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGG 1947
Db 3079 TCTCTGCGTGGATATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGG 3138
QY 1948 GAGGATCCGTGTCTCTGCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTGGA 2007
Db 3139 GAGGATCCGTGTCTCTGCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTGGA 3198
QY 2008 AGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTC AACAGGATTTTGTGACCAGCG 2067
Db 3199 AGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTC AACAGGATTTTGTGACCAGCG 3258

QY 2068 CAGGCTGGGCTCCTTCTGTCATGATTCTATCCAA 2101
Db 3259 CAGGCTGGGCTCCTTCTGTCATGATTCTATCCAA 3292
RESULT 4
AR304538
LOCUS AR304538 5952 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 1 from patent US 5544786.
ACCESSION AR304538
VERSION AR304538.1 GI:31693691
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5952)
AUTHORS Xiao,X. and Liu,P.X.
TITLE Method and vector for producing and transferring trans-spliced peptides
JOURNAL Patent: US 5544786-A 1 08-APR-2003;
FEATURES Location/Qualifiers
source 1. .5952 /organism="unknown"
BASE COUNT 1860 a 1344 c 1410 g 1338 t
ORIGIN
Query Match 63.2%; Score 1328.8; DB 6; Length 5952;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 762 GGGTTCTTTTACAAGACAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAAC 821
Db 3284 GGTCCCATTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAAC 3343
QY 822 TTCTGGTGTGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCG 881
Db 3344 TTCTGGTGTGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCG 3403
QY 882 ACTTTCAGCAGTTTCAAGACGAAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAA 941
Db 3404 ACTTTCAGCAGTTTCAAGACGAAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAA 3463
QY 942 CTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGCAGAGCAGC 1001
Db 3464 CTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGCAGAGCAGC 3523
QY 1002 CTTTGGAGGACTAGAGAAACTCTACAGGAGCCCCAGAGAGCTGCCTCCTGAGGAGAG 1061
Db 3524 CTTTGGAGGACTAGAGAAACTCTACAGGAGCCCCAGAGAGCTGCCTCCTGAGGAGAG 3583
QY 1062 CCCAGAAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAA 1121
Db 3584 CCCAGAAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAA 3643
QY 1122 AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCC 1181
Db 3644 AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCC 3703
QY 1182 AGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCA 1241
Db 3704 AGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCA 3763
QY 1242 AGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGA 1301
Db 3764 AGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGA 3823
QY 1302 AAGTCAAGGCACTTCGAGGAGAAATTGGCCTCTGAAAGAGAACGTCAGCCACGTCATG 1361
Db 3824 AAGTCAAGGCACTTCGAGGAGAAATTGGCCTCTGAAAGAGAACGTCAGCCACGTCATG 3883
QY 1362 ACCTTGCTCGCCAGCTTACCACCTTTGGGATTCAGCTCTCACCGTATAACCTCAGCACTC 1421
Db 3884 ACCTTGCTCGCCAGCTTACCACCTTTGGGATTCAGCTCTCACCGTATAACCTCAGCACTC 3943

```
Qy 1422 TGAAGACCTGAACACAGATGGAAGCTTCTGAGGTGGCCGTCGAGACCGAGTCAGGC 1481
Db 3944 TGAAGACCTGAACACAGATGGAAGCTTCTGAGGTGGCCGTCGAGACCGAGTCAGGC 4003
Qy 1482 AGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTG 1541
Db 4004 AGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTG 4063
Qy 1542 TCCAGGGTCCCTGGGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACG 1601
Db 4064 TCCAGGGTCCCTGGGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACG 4123
Qy 1602 AGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTTAGCTG 1661
Db 4124 AGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTTAGCTG 4183
Qy 1662 ACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 1721
Db 4184 ACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 4243
Qy 1722 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGACCAGCACA 1781
Db 4244 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGACCAGCACA 4303
Qy 1782 ACCTCAAGCAAAATGACCGCCCATGGATATCTCTGCAGATTATTAATTGTTTGACCACTA 1841
Db 4304 ACCTCAAGCAAAATGACCGCCCATGGATATCTCTGCAGATTATTAATTGTTTGACCACTA 4363
Qy 1842 TTTATGACCGCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTGCGTGGATA 1901
Db 4364 TTTATGACCGCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTGCGTGGATA 4423
Qy 1902 TGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCC 1961
Db 4424 TGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCC 4483
Qy 1962 TGTCTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGAT 2021
Db 4484 TGTCTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGAT 4543
Qy 2022 ACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCGGACGGCTGGGCTTCC 2081
Db 4544 ACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCGGACGGCTGGGCTTCC 4603
Qy 2082 TTCTGCAATGATTCTATCCAA 2101
Db 4604 TTCTGCAATGATTCTATCCAA 4623

RESULT 5
AX114289
LOCUS AX114289 5952 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 1 from Patent WO0129243.
ACCESSION AX114289
VERSION AX114289.1 GI:14031259
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Paul,X.L. and Xiao,X.
TITLE Method and vector for producing and transferring trans -spliced
peptides
JOURNAL Patent: WO 0129243-A 1 26-APR-2001;
DALHOUSIE UNIVERSITY (CA) ; UNIV. OF PITTSBURGH OF THE COMMONWEALTH
SYSTEM OF HIGHER EDUCATION (US)
FEATURES
source Location/Qualifiers
1. .5952
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
```

```
misc_feature 2897..2898
misc_feature /note="S4 junction site"
misc_feature 3198..3199
misc_feature /note="S2 junction site"
BASE COUNT 1860 a 1344 c 1410 g 1338 t
ORIGIN
Query Match 63.2%; Score 1328.8; DB 6; Length 5952;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 762 GGGTTCTTTTACAAGACAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTTCTGCAGGAAC 821
Db 3284 GGTCCCATTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTTCTGCAGGAAC 3343
Qy 822 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCG 881
Db 3344 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCG 3403
Qy 882 ACTTTCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAA 941
Db 3404 ACTTTCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAA 3463
Qy 942 CTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 1001
Db 3464 CTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 3523
Qy 1002 CTTTGAAGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAG 1061
Db 3524 CTTTGAAGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAG 3583
Qy 1062 CCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAA 1121
Db 3584 CCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAA 3643
Qy 1122 AATTGAACCTGCACCTCGCTGACTGGCAGAGAGAAATAGATGAGACCCTTGAAGACTCC 1181
Db 3644 AATTGAACCTGCACCTCGCTGACTGGCAGAGAGAAATAGATGAGACCCTTGAAGACTCC 3703
Qy 1182 AGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCA 1241
Db 3704 AGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCA 3763
Qy 1242 AGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGA 1301
Db 3764 AGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGA 3823
Qy 1302 AAGTCAAGGCACCTTCGAGGAGAGAAATTGCGCTCTGAAAGAGAACGTCGAGCCAGTCAATG 1361
Db 3824 AAGTCAAGGCACCTTCGAGGAGAGAAATTGCGCTCTGAAAGAGAACGTCGAGCCAGTCAATG 3883
Qy 1362 ACCTTGCTGCCAGCTTACACATTTGGGCACTTCAGCTCTCACCGTATAACCTCAGCACTC 1421
Db 3884 ACCTTGCTGCCAGCTTACACATTTGGGCACTTCAGCTCTCACCGTATAACCTCAGCACTC 3943
Qy 1422 TGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGC 1481
Db 3944 TGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGC 4003
Qy 1482 AGCTGCATGAAGCCCAACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTG 1541
Db 4004 AGCTGCATGAAGCCCAACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTG 4063
Qy 1542 TCCAGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACG 1601
Db 4064 TCCAGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACG 4123
Qy 1602 AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGACTCTACCACTTTTAGCTG 1661
Db 4124 AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGACTCTACCACTTTTAGCTG 4183
Qy 1662 ACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 1721
```


| | | | |
|---|------|--|------|
| Db | 4184 | ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA | 4243 |
| QY | 1722 | AGGCCCTTTGCTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCTTGAGACCAGACA | 1781 |
| Db | 4244 | AGGCCCTTTGCTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCTTGAGACCAGACA | 4303 |
| QY | 1782 | ACCTCAAGCAAAATGACCGCCATGGATATCCTGCAGATTATAATTGTTTGACCACTA | 1841 |
| Db | 4304 | ACCTCAAGCAAAATGACCGCCATGGATATCCTGCAGATTATAATTGTTTGACCACTA | 4363 |
| QY | 1842 | TTTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCCTCTCTGCGTGGATA | 1901 |
| Db | 4364 | TTTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCCTCTCTGCGTGGATA | 4423 |
| QY | 1902 | TGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCC | 1961 |
| Db | 4424 | TGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCC | 4483 |
| QY | 1962 | TGTCTTTTAAACTGGCATATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGAT | 2021 |
| Db | 4484 | TGTCTTTTAAACTGGCATATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGAT | 4543 |
| QY | 2022 | ACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTACCAAGCGCAGGCTGGCCCTCC | 2081 |
| Db | 4544 | ACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTACCAAGCGCAGGCTGGCCCTCC | 4603 |
| QY | 2082 | TTCTGCATGATTCTATCCAA | 2101 |
| Db | 4604 | TTCTGCATGATTCTATCCAA | 4623 |
| RESULT 6 | | | |
| AX538622 | | | |
| LOCUS | | | |
| AX538622 | | | |
| DEFINITION | | | |
| Sequence 42 from Patent WO0229056. | | | |
| ACCESSION | | | |
| AX538622 | | | |
| VERSION | | | |
| AX538622.1 GI:25271171 | | | |
| KEYWORDS | | | |
| synthetic construct | | | |
| SOURCE | | | |
| ORGANISM | | | |
| artificial sequences. | | | |
| REFERENCE | | | |
| 1 | | | |
| AUTHORS | | | |
| Chamberlain,J.S. and Harper,S.Q. | | | |
| TITLE | | | |
| Mini-dystrophin nucleic acid and peptide sequences | | | |
| JOURNAL | | | |
| Patent: WO 0229056-A 42 11-APR-2002; | | | |
| THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US) | | | |
| FEATURES | | | |
| source | | | |
| 1. .8689 | | | |
| /organism="synthetic construct" | | | |
| /mol_type="genomic DNA" | | | |
| /db_xref="taxon:32630" | | | |
| /note="Synthetic" | | | |
| BASE COUNT | | | |
| 2721 a 1804 c 1861 g 2303 t | | | |
| ORIGIN | | | |
| Query Match 63.2%; Score 1328.8; DB 6; Length 8689; | | | |
| Best Local Similarity 99.5%; Pred. No. 0; | | | |
| Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0; | | | |
| QY | 762 | GGGTTCTTTTACAAGACAGTTCTGACCAAGTGAAGCGTCTGACCTTTCTCTGCAAGAAC | 821 |
| Db | 3330 | GGTCCCATTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACTTTCTCTGCAAGAAC | 3389 |
| QY | 822 | TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCG | 881 |
| Db | 3390 | TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCG | 3449 |
| QY | 882 | ACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAA | 941 |
| Db | 3450 | ACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAA | 3509 |
| QY | 942 | CTAAAGAACCTGTAATCATGAGTACTCTTCAGACTGTACGAATATTTCTGACAGAGCAGC | 1001 |

| | | | |
|----|------|---|------|
| Db | 3510 | CTAAAGAACCTGTATAATCATGAGTACTTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC | 3569 |
| QY | 1002 | CTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCCAGAGAGCTGCCTCCTTGAGGAGAGAG | 1061 |
| Db | 3570 | CTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCCAGAGAGCTGCCTCCTTGAGGAGAGAG | 3629 |
| QY | 1062 | CCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAA | 1121 |
| Db | 3630 | CCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAA | 3689 |
| QY | 1122 | AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAAGACTCC | 1181 |
| Db | 3690 | AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAAGACTCC | 3749 |
| QY | 1182 | AGGAACCTTCAAGAGGCCACCGGATGAGCTGGAACCTCAAGCTGCGCCAAGCTGAGGTGATCA | 1241 |
| Db | 3750 | AGGAACCTTCAAGAGGCCACCGGATGAGCTGGAACCTCAAGCTGCGCCAAGCTGAGGTGATCA | 3809 |
| QY | 1242 | AGGATCCTTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCTCAAGATCACCTCGAGA | 1301 |
| Db | 3810 | AGGATCCTTGGCAGCCCGTGGCGGATCTCCTCATTGACTCTCTCCTCAAGATCACCTCGAGA | 3869 |
| QY | 1302 | AAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGCTGAGCCACGTCAAATG | 1361 |
| Db | 3870 | AAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGCTGAGCCACGTCAAATG | 3929 |
| QY | 1362 | ACCTTGCTCGCCAGCTTACCACATTTGGGCATTCAGCTCTCACCGTATAAACCTCAGCACTC | 1421 |
| Db | 3930 | ACCTTGCTCGCCAGCTTACCACATTTGGGCATTCAGCTCTCACCGTATAAACCTCAGCACTC | 3989 |
| QY | 1422 | TGGAAGACCTGAAACACCAAGATGGAAGCTTCTGCAGGTGGCCCTCGAGGACCGAGTCAGGC | 1481 |
| Db | 3990 | TGGAAGACCTGAAACACCAAGATGGAAGCTTCTGCAGGTGGCCCTCGAGGACCGAGTCAGGC | 4049 |
| QY | 1482 | AGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGCACTTCTTTCCACGCTCG | 1541 |
| Db | 4050 | AGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGCACTTCTTTCCACGCTCG | 4109 |
| QY | 1542 | TCCAGGTCCTTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACG | 1601 |
| Db | 4110 | TCCAGGTCCTTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACG | 4169 |
| QY | 1602 | AGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTG | 1661 |
| Db | 4170 | AGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTG | 4229 |
| QY | 1662 | ACCTGAATAATGTGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA | 1721 |
| Db | 4230 | ACCTGAATAATGTGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA | 4289 |
| QY | 1722 | AGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCTTGGACCAGACA | 1781 |
| Db | 4290 | AGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCTTGGACCAGACA | 4349 |
| QY | 1782 | ACCTCAAGCAAAAATGACCGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACCTA | 1841 |
| Db | 4350 | ACCTCAAGCAAAAATGACCGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACCTA | 4409 |
| QY | 1842 | TTTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCCTCTCTCGTGGATA | 1901 |
| Db | 4410 | TTTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCCTCTCTCGTGGATA | 4469 |
| QY | 1902 | TGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCC | 1961 |
| Db | 4470 | TGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCC | 4529 |
| QY | 1962 | TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGAT | 2021 |
| Db | 4530 | TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGAT | 4589 |
| QY | 2022 | ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTACCAAGCGCAGGCTGGGCCTCC | 2081 |
| Db | 4590 | ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTACCAAGCGCAGGCTGGGCCTCC | 4649 |

Db 9628 TGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTTGTC 9687
Qy 1962 TGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAACTACAGAT 2021
Db 9688 TGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAACTACAGAT 9747
Qy 2022 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCCTCC 2081
Db 9748 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCCTCC 9807
Qy 2082 TTCTGCATGATTCTATCCAA 2101
Db 9808 TTCTGCATGATTCTATCCAA 9827
RESULT 10
AX409637
LOCUS AX409637 13957 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 2284 from Patent WO0229103.
ACCESSION AX409637
VERSION AX409637.1 GI:21442342
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2284 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
source
1. .13957
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. M18533"
BASE COUNT 4602 a 2781 c 3122 g 3452 t
ORIGIN
Query Match 63.2%; Score 1328.8; DB 6; Length 13957;
Best local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 762 GGGTTCTTTTACAAGACAGTTCTGACCAGTGAAGCGTCTGCACCTTTTCTCTGCAGGAAC 821
Db 8598 GGTCCCATTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTTCTCTGCAGGAAC 8657
Qy 822 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGAGGCG 881
Db 8658 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGAGGCG 8717
Qy 882 ACTTCCAGCAGTTTCAAGACGAAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAA 941
Db 8718 ACTTCCAGCAGTTTCAAGACGAAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAA 8777
Qy 942 CTAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 1001
Db 8778 CTAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 8837
Qy 1002 CTTTGGAAAGGACTAGAGAAACTCTACAGAGAGCCCGCAGAGAGTGCCTCCTGAGGAGAGAG 1061
Db 8838 CTTTGGAAAGGACTAGAGAAACTCTACAGAGAGCCCGCAGAGAGTGCCTCCTGAGGAGAGAG 8897
Qy 1062 CCCAGAAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAA 1121
Db 8898 CCCAGAAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAA 8957
Qy 1122 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCC 1181
Db 8958 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCC 9017

Qy 1182 AGGAACCTTCAAGAGGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCA 1241
Db 9018 AGGAACCTTCAAGAGGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCA 9077
Qy 1242 AGGATCCTTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTTCCAAGATCACCTCGAGA 1301
Db 9078 AGGATCCTTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTTCCAAGATCACCTCGAGA 9137
Qy 1302 AAGTCAAGGCACCTTCGAGGAGAAATTTGCGCCTCTGAAAGAGAAAGCTGAGGCCACGTCAATG 1361
Db 9138 AAGTCAAGGCACCTTCGAGGAGAAATTTGCGCCTCTGAAAGAGAAAGCTGAGGCCACGTCAATG 9197
Qy 1362 ACCTTGCTGCCAGCTTACCACTTTTGGGCATTCAGCTCTCACCGTATAAAGCTCAGCACTC 1421
Db 9198 ACCTTGCTGCCAGCTTACCACTTTTGGGCATTCAGCTCTCACCGTATAAAGCTCAGCACTC 9257
Qy 1422 TGGAAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTGGCCGTTCGAGGACCGAGTCAGGC 1481
Db 9258 TGGAAAGACCTGAACACCCAGATGGAAGCTTCTGCAGGTGGCCGTTCGAGGACCGAGTCAGGC 9317
Qy 1482 AGCTGCATGAAGCCACACAGGAGCTTTGGTCCAGCATCTCAGCACTCTTTTCCACGCTGTG 1541
Db 9318 AGCTGCATGAAGCCACACAGGAGCTTTGGTCCAGCATCTCAGCACTCTTTTCCACGCTGTG 9377
Qy 1542 TCCAGGTCCTCGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACACG 1601
Db 9378 TCCAGGTCCTCGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACACG 9437
Qy 1602 AGACTCAAAACAACTTGTCTGGGACCATCCCAAATGACAGAGCTCTACCACTCTTTAGCTG 1661
Db 9438 AGACTCAAAACAACTTGTCTGGGACCATCCCAAATGACAGAGCTCTACCACTCTTTAGCTG 9497
Qy 1662 ACCTGAATAATGTCTCAGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 1721
Db 9498 ACCTGAATAATGTCTCAGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 9557
Qy 1722 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTCTCAGCTGCATGTGATCCCTTGACCAGCACA 1781
Db 9558 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTCTCAGCTGCATGTGATCCCTTGACCAGCACA 9617
Qy 1782 ACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTA 1841
Db 9618 ACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTA 9677
Qy 1842 TTTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAAGTCCCTCTCTCGCTGGATA 1901
Db 9678 TTTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAAGTCCCTCTCTCGCTGGATA 9737
Qy 1902 TGTGCTGAACCTGGCTGCTGAATGTTTTATGATACGGGACGAAACAGGAGGATCCGTGTCC 1961
Db 9738 TGTGCTGAACCTGGCTGCTGAATGTTTTATGATACGGGACGAAACAGGAGGATCCGTGTCC 9797
Qy 1962 TGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGAT 2021
Db 9798 TGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGAT 9857
Qy 2022 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCTCC 2081
Db 9858 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCTCC 9917
Qy 2082 TTCTGCATGATTCTATCCAA 2101
Db 9918 TTCTGCATGATTCTATCCAA 9937
RESULT 11
AX538581
LOCUS AX538581 13957 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 1 from Patent WO0229056.
ACCESSION AX538581
VERSION AX538581.1 GI:25271086
KEYWORDS Homo sapiens (human)
SOURCE

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE Chamberlain, J.S. and Harper, S.Q.
AUTHORS Mini-dystrophin nucleic acid and peptide sequences
TITLE Patent: WO 0229056-A 1 11-APR-2002;
JOURNAL THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES Location/Qualifiers
source
1..13957
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 4602 a 2781 c 3122 g 3452 t
ORIGIN
Query Match 63.2%; Score 1328.8; DB 6; Length 13957;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 762 GGGTCTCTTTTACAGACAGATTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAAC 821
DB 8598 GGTCCCATTTGGAAGCCAGTTCTTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAAC 8657
QY 822 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAAGCCGGCAGGCACCTATTGGAGCG 881
DB 8658 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAAGCCGGCAGGCACCTATTGGAGCG 8717
QY 882 ACTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAAATTGAAA 941
DB 8718 ACTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAAATTGAAA 8777
QY 942 CTAAGAACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGC 1001
DB 8778 CTAAGAACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGC 8837
QY 1002 CTTTGGAGGACTAGAGAAACTCTACCAGAGCCCGAGAGAGCTGCCTCCTGAGGAGAGAG 1061
DB 8838 CTTTGGAGGACTAGAGAAACTCTACCAGAGCCCGAGAGAGCTGCCTCCTGAGGAGAGAG 8897
QY 1062 CCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAA 1121
DB 8898 CCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAA 8957
QY 1122 AATTGAACCTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCC 1181
DB 8958 AATTGAACCTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCC 9017
QY 1182 AGGAACCTTCAAGAGGCGCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCA 1241
DB 9018 AGGAACCTTCAAGAGGCGCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCA 9077
QY 1242 AGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGA 1301
DB 9078 AGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGA 9137
QY 1302 AAGTCAAGGCACCTTCGAGGAGAAATTGGCGCTCTGAAAGAGAACGTCGAGCCACGTCAATG 1361
DB 9138 AAGTCAAGGCACCTTCGAGGAGAAATTGGCGCTCTGAAAGAGAACGTCGAGCCACGTCAATG 9197
QY 1362 ACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTC 1421
DB 9198 ACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTC 9257
QY 1422 TGGAGACCTGAAACACCAAGATGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGC 1481
DB 9258 TGGAGACCTGAAACACCAAGATGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGC 9317
QY 1482 AGCTGCATGAAGCCCAAGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTG 1541
DB 9318 AGCTGCATGAAGCCCAAGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTG 9377
QY 1542 TCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCACG 1601

Db 9378 TCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACG 9437
QY 1602 AGACTCAAAACAACTTGTCTGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTG 1661
DB 9438 AGACTCAAAACAACTTGTCTGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTG 9497
QY 1662 ACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 1721
DB 9498 ACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 9557
QY 1722 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTTCAGCTGCATGTGATGCCTTGGACCAACACA 1781
DB 9558 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTTCAGCTGCATGTGATGCCTTGGACCAACACA 9617
QY 1782 ACCTCAAGCAAAATGACCAAGAGCAGCAACAATTTGGTCAACGTCCTCTCTGCGTGGATA 1841
DB 9618 ACCTCAAGCAAAATGACCAAGAGCAGCAACAATTTGGTCAACGTCCTCTCTGCGTGGATA 9677
QY 1842 TTTATGACCGCCTGGAGCAAGAGCAGCAACAATTTGGTCAACGTCCTCTCTGCGTGGATA 1901
DB 9678 TTTATGACCGCCTGGAGCAAGAGCAGCAACAATTTGGTCAACGTCCTCTCTGCGTGGATA 9737
QY 1902 TGTGCTGAACTGGCTGCTGATGATTTTATGATACCGGACGAGGAGGATCCGCTGTC 1961
DB 9738 TGTGCTGAACTGGCTGCTGATGATTTTATGATACCGGACGAGGAGGATCCGCTGTC 9797
QY 1962 TGTCTTTTAAACTGGCATCAATTTCCCTGTGTAAAGCAGATTTTGGAGACAAGTACAGAT 2021
DB 9798 TGTCTTTTAAACTGGCATCAATTTCCCTGTGTAAAGCAGATTTTGGAGACAAGTACAGAT 9857
QY 2022 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGGCTGGGCTCC 2081
DB 9858 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGGCTGGGCTCC 9917
QY 2082 TTCTGCATGATTTCTATCCAA 2101
DB 9918 TTCTGCATGATTTCTATCCAA 9937
RESULT 12
HUMDYS
LOCUS HUMDYS 13957 bp mRNA linear PRI 25-MAY-2000
DEFINITION Homo sapiens dystrophin (DMD) mRNA, complete cds.
ACCESSION M18533 M17154 M18026 M20250
VERSION M18533.1 GI:181856
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1699)
AUTHORS Koenig, M., Hoffman, E.P., Bertelson, C.J., Monaco, A.P., Feener, C. and Kunkel, L.M.
TITLE Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals
JOURNAL Cell 50 (3), 509-517 (1987)
MEDLINE 87273512
PUBMED 3607877
REFERENCE 2 (bases 1678 to 3830)
AUTHORS Hoffman, E.P., Monaco, A.P., Feener, C.C. and Kunkel, L.M.
TITLE Conservation of the Duchenne muscular dystrophy gene in mice and humans
JOURNAL Science 238 (4825), 347-350 (1987)
MEDLINE 88018015
PUBMED 3659917
REFERENCE 3 (bases 1 to 13957)
AUTHORS Koenig, M., Monaco, A.P. and Kunkel, L.M.
TITLE The complete sequence of dystrophin predicts a rod-shaped cytoskeletal protein
JOURNAL Cell 53 (2), 219-226 (1988)
MEDLINE 88194521

3282674

PUBMED

COMMENT

On May 25, 2000 this sequence version replaced gi:340693.
Draft entry and computer-readable sequence kindly provided by
M.Koenig, 01-APR-1988 The severity of muscular dystrophy is
determined by the size of the deleted DNA segment. Deletions found
in different patients were from positions 302-2200, 473-1168,
1691-1810, and 1169-3011.

FEATURES

Location/Qualifiers

1..13957
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp21.3-p21.1"
/tissue_type="muscle"
/dev_stage="fetus"
/note="G00-119-850"

gene

CDS

1..13957
/gene="DMD"
209..11266
/gene="DMD"
/codon_start=1

/product="dystrophin"
/protein_id="AAA53189.1"
/db_xref="GI:181857"

/translation="MLWEEVEDCYEREDVQKFTFTKWNAQFSKFGKHQHIENLFSDL
QDGRRLDLDLEGLTGQKLPKEKSTVRHALNNVNKALRVLQNNVLDVNLVIGSTDIVDG
NHLKTLGLIWNIIILHWQVNMKNIMAGLOCTNSEKILLSWVRQSTRYPQVNVINFT
TWSVDGLNALNLIHSHRDLFDWNSVVCQASATQLEHAFNIARYQLGIEKLDPEDV
DTTPDKKSILMYITSLFQVLPQQVSIEAIEQEVEMLPKPKVTKEEHPQLHQMHSQ
QITVSLAQGYERTSSPKPRFKSYAYTQAAVYVTSDFTRSPFPSSHLEAPEDKSPGSSL
MESEVNLDRYQTALREVLSWLLSAEDTLQAQGEISNDVEVVKQOQFHTHEGYMMDLTAH
QCRVGNILQGLSKLIGTKLSEDEETEVEQOMNLLNSRWECLRVASKEQSNLHRVLM
DLQNGKIKELNDLTKTEERTKMEEEPLGPDLEDLKRQVQKHVQLQEDLEQEQVRVN
SLTHMVVVDESSGHDATAALEEQKLVGLDRWANI CRWTEDRWVLLQDILLKWQRLTE
EQCLFSAWLSEKEDAVNKIHTTGFKDQNEMLSSIQKLAVLKADLEKKQSMGKLYSLK
QDLSLTNKSQVTQKTEAWLNDNFARCDNLVQKLEKSTAQISQAVITTPQSLTQTTVM
ETVTVTTRTEQILVKHAQSELP PPPQKKRQITVDSEIRKRLDVIDITELHSWITRSEA
VLQSPFEAIFRKEGNSDLKEKVNAIEREKAERKLDQASRAQALVEQMNEGVNA
DSIKQASEQLNSRWIEFCQLLSERLNLWLEYQNNIIAFYNQLQOLEQMTTAEANWLXIQ
TTPSEPTAIKSKLKI KDENVRLSGLQPOQIERLKIQSI ALKEKGQGMFLDADFVAF
TNHFQVSDVQAREKELQITFDLTPPMRYQETWSAIRTWQVQSETKLSIPQLSVTDY
EIMEORLGELOALQSSLOEQSGLYLSTTVKESKAPSEISRKYQSFEFEEIEGRWK
KLSSQLVEHCQKLEEQMNLKRIQNHIOQLKMAEVDVFLKEEWPALGDSILKKQL
KOCRLIVSDIOTIQPSLNSVNEGQKI KNEAPEPASRLTEKELNTQWDMCCQVY
ARKEALKGGLKTVSLQDLSMHEWMTQAEVEYLERDFEYKTPDELQKAVEEMKRAK
EEAQOKEAKVLLTESVNSVIAQAPPVQAEALKKLELTITNYQWLCTRLNGCKKLTLE
EVMACHHELLSYLEKANKWNEVEFKLKTENIPGAAEIESEVLDLENLMRHSNDNP
NQIRILAQTLTDGGVMDLINELETFNSRRREHAEVRRQKLEQSIQSAQETESK
LHLIQESLTFIDKQLAAYTADKVDAAQMPQEAQKIQSDLSHSIEISLEEMKHNQKEA
AQVLSQIDVAQKQLQDVSMKFRLLFKQPFANFELRLDEVKMHLPLPALETQKSV
QEVVQSLNHCNVNLYKSLSEYKSEVEMVIKTGRQIVQKKQOTENPKELDERVTALKLHY
NELGAKVTERKQOLEKCLKLSRKMVKEMNVLTEWLAATDMELTKRS AVEGMPSLNDS
VANGKATQKEIEKQKVLKSTIEVGEALKTVLGKXETLVEDKLSSLNSNWLAVTSRAE
EWNLLLEYQKHMETFDQNVDTHTKI IQADTLDESEKKKPPQKEDVLKRLKAELEND
IRPKVDSTRDQAAANLMA NRGDHCKLVEPOISELNRHFAAISHR IKTGKASIPLKELE
QFNSDIQKLEPLEAEIQQVNLKEEDFNKDMNEDNEGTVKELLQRGDNLQORITDER
KREEIKIQQLLOTKHNALDLRSQRRKALEISHQWYQYKQADDDLLKCLDDIEKKL
ASLPEPDERKIKEIDRELQKKKEELNVRROAEGISEDGAAMAVEPTQILSKWRE
IESKFAQFRLNFAQIHTVREETMMVMTEDMPLEISYVPSYLTETI THVSQALLEVEQ
LNLAPDLCAKDFEDLFQKEESLKNIKDSLQOSSGRDIIHKKTAALQSATPVERVKL
OEALSOLDPQWEKVNKMYKDRQGRFDRSVEKRRRFDYDIKIFNQLWTEAEQFLRKTQI
PENWEHAKYKWLKELQDGIGCGRQTVFRTLNAATGEEIIQQSSKTASILEKGLSLNL
RWQEVCKLSDRKRLLEQXNIISEFORDNLNEFVWLEADNIIASIPLEPGKEQQLKE
KLEQVKLVEELPLRQGLKQNETGGPVLVSAPISEEQDKLENLKTQNLQWIKVS
RALPEKQGEIEAQIKDLGLEKLELDEEQNLHLLWLSPIRNOLEIYNQNOEGPFD
VQETETAVQAKQPDVEEILSKQHLKYEKPATQPVKRLKLEDLSSEWKNVRLQELRA
KQPDLAGLTTIGASPTQTVTLTPVVTKETATISLEMPSSLMLEVPALADFNRAWT
ELTDWLSLDDQVIKSQRVMVGLEDINEMIIKQKATMDLEQRRRPQLEELITAAQNLK
NKTNSQEARTIITDRIERI QNQWDEVQEHQLNRRQQLNEMLKDSTQWLEAKEBAEQVL
GOARAKLESWKEGPTYVDIAIQKXITETKQALAKOLRWQTNVDVANDLALKLLRDYSAD
DTRKVTMTENINASWRSIHKRVSEREALEETHRLIQFPFLDKFLAWLTAETTA
NVLQDATRKERLLEDSSKGVKELNKQWQDLQGEIEAHTDVYHNLDNSQKILRSLEGSD

| | | | |
|--|------|--|------|
| Db | 9318 | AGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTG | 9377 |
| QY | 1542 | TCCAGGFTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACG | 1601 |
| Db | 9378 | TCCAGGFTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACG | 9437 |
| QY | 1602 | AGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTG | 1661 |
| Db | 9438 | AGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTG | 9497 |
| QY | 1662 | ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA | 1721 |
| Db | 9498 | ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA | 9557 |
| QY | 1722 | AGGCCCTTTGCTTGGATCTCTTTGAGCCTGTCAGCTGCATGTGATGCCCTTGACCCAGACA | 1781 |
| Db | 9558 | AGGCCCTTTGCTTGGATCTCTTTGAGCCTGTCAGCTGCATGTGATGCCCTTGACCCAGACA | 9617 |
| QY | 1782 | ACCTCAAGCAAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCACTA | 1841 |
| Db | 9618 | ACCTCAAGCAAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCACTA | 9677 |
| QY | 1842 | TTTATGACCCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCCTCTCTGCGTGGATA | 1901 |
| Db | 9678 | TTTATGACCCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCCTCTCTGCGTGGATA | 9737 |
| QY | 1902 | TGTGCTGAATGGCTGCTGTAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCC | 1961 |
| Db | 9738 | TGTGCTGAATGGCTGCTGTAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCC | 9797 |
| QY | 1962 | TGTCTTTTAAACTGGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGAT | 2021 |
| Db | 9798 | TGTCTTTTAAACTGGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGAT | 9857 |
| QY | 2022 | ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAAGGAGGATCCGTGTCC | 2081 |
| Db | 9858 | ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAAGGAGGATCCGTGTCC | 9917 |
| QY | 2082 | TTCTGCATGATTCTATCCAA 2101 | |
| Db | 9918 | TTCTGCATGATTCTATCCAA 9937 | |
| RESULT 13 | | | |
| AR220819 | | | |
| LOCUS | | | |
| DEFINITION | | | |
| SEQUENCE 60 from patent US 6426186. | | | |
| AR220819 | | | |
| AR220819.1 | | | |
| GI:23327696 | | | |
| KEYWORDS | | | |
| SOURCE | | | |
| ORGANISM | | | |
| Unclassified. | | | |
| REFERENCE | | | |
| 1 (bases 1 to 13977) | | | |
| JONES, K.A., Volkmuth, W. and Walker, M.G. | | | |
| AUTHORS | | | |
| TITLE | | | |
| Bone remodeling genes | | | |
| JOURNAL | | | |
| Patent: US 6426186-A 60 30-JUL-2002; | | | |
| FEATURES | | | |
| Location/Qualifiers | | | |
| 1..13977 | | | |
| /organism="unknown" | | | |
| BASE COUNT | | | |
| 4596 a 2765 c 3120 g 3453 t 43 others | | | |
| ORIGIN | | | |
| Query Match | | | |
| Best Local Similarity 62.7%; Score 1317.8; DB 6; Length 13977; | | | |
| Matches 1333; Conservative 99.4%; Pred. No. 0; | | | |
| 0; Mismatches 7; Indels 1; Gaps 1; | | | |
| QY | 762 | GGGTTCTTTTACAAGACAGTCTTGACCAAGTGGGAAGCGTGTGCACCTTTCTCTGCAGGAAC | 821 |
| Db | 8598 | GGTCCCATTTGGAAGCCAGTTCTGACCAAGTGGGAAGCGTGTGCACCTTTCTCTGCAGGAAC | 8657 |
| QY | 822 | TTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCG | 881 |

| | | | |
|----|------|--|------|
| Db | 8658 | TTCTGCTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCG | 8717 |
| QY | 882 | ACTTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAAATTGAAAA | 941 |
| Db | 8718 | ACTTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAAATTGAAAA | 8777 |
| QY | 942 | CTAAAGAACCTGTATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGC | 1001 |
| Db | 8778 | CTAAAGAACCTGTATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGC | 8837 |
| QY | 1002 | CTTTGGAAGGACTAGAGAACTCTACCAGGAGCCCAAGAGAGCTGCCTCCTGAGGAGAGAG | 1061 |
| Db | 8838 | CTTTGGAAGGACTAGAGAACTCTACCAGGAGCCCAAGAGAGCTGCCTCCTGAGGAGAGAG | 8897 |
| QY | 1062 | CCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAA | 1121 |
| Db | 8898 | CCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAA | 8957 |
| QY | 1122 | AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCC | 1181 |
| Db | 8958 | AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCC | 9017 |
| QY | 1182 | AGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCA | 1241 |
| Db | 9018 | AGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCA | 9077 |
| QY | 1242 | AGGGATCCTGGCAGCCCCGTGGCGATCTCTCATTTGACTCTCTCCAAGATCACCTCGAGA | 1301 |
| Db | 9078 | AGGGATCCTGGCAGCCCCGTGGCGATCTCTCATTTGACTCTCTCCAAGATCACCTCGAGA | 9137 |
| QY | 1302 | AAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAAGAGAACGTCGAGCCACGTC | 1361 |
| Db | 9138 | AAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAAGAGAACGTCGAGCCACGTC | 9197 |
| QY | 1362 | ACCTTGCTCGCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTC | 1421 |
| Db | 9198 | ACCTTGCTCGCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTC | 9257 |
| QY | 1422 | TGGAAGACCTTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGC | 1481 |
| Db | 9258 | TGGAAGACCTTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGC | 9317 |
| QY | 1482 | AGCTGCATGAAGCCACAGGGACCTTTGGTCCAGCATCTCAGCACTTTCTTCCACGCTCTG | 1541 |
| Db | 9318 | AGCTGCATGAAGCCACAGGGACCTTTGGTCCAGCATCTCAGCACTTTCTTCCACGCTCTG | 9377 |
| QY | 1542 | TCCAGGGTCCCTGGGAGAGAGCCATCTCGSCAAAACAAAGTGCCCTACTATATCAACCACG | 1601 |
| Db | 9378 | TCCAGGGTCCCTGGGAGAGAGCCATCTCGSCAAAACAAAGTGCCCTACTATATCAACCACG | 9437 |
| QY | 1602 | AGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTTAGCTG | 1661 |
| Db | 9438 | AGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTTAGCTG | 9497 |
| QY | 1662 | ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA | 1721 |
| Db | 9498 | ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA | 9557 |
| QY | 1722 | AGGCCCTTTGCTTGGATCTCTTTGAGCCTGTGACCTGTGATGCTGATGCCCTTGACCAGCACA | 1781 |
| Db | 9558 | AGGCCCTTTGCTTGGATCTCTTTGAGCCTGTGACCTGTGATGCTGATGCCCTTGACCAGCACA | 9617 |
| QY | 1782 | ACCTCAAGCAAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCACTA | 1841 |
| Db | 9618 | ACCTCAAGCAAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCACTA | 9677 |
| QY | 1842 | TTTATGACCCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCCTCTCTGCGTGGATA | 1901 |
| Db | 9678 | TTTATGACCCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCCTCTCTGCGTGGATA | 9737 |
| QY | 1902 | TGTGCTGAATGGCTGCTGTAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCC | 1961 |
| Db | 9738 | TGTGCTGAATGGCTGCTGTAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCC | 9797 |

| | | | |
|-----------------------|--|--|-----------------------------------|
| QY | 1962 | TGTCCTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGAAGACAAGTACAGAT | 2021 |
| Db | 9798 | TGTCCTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGAAGACAAGTACAGAT | 9857 |
| QY | 2022 | ACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGCGCAGGCT-GGGCCTC | 2080 |
| Db | 9858 | ACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGGGCCTC | 9917 |
| QY | 2081 | CTTCTGCATGATCTATCCAA | 2101 |
| Db | 9918 | CTTCTGCATGATCTATCCAA | 9938 |
| RESULT 14 | | | |
| E30219 | | 4402 bp | DNA |
| LOCUS | E30219 | Shortened dystrophin. | linear |
| DEFINITION | E30219 | Shortened dystrophin. | PAT 18-JUN-2001 |
| ACCESSION | E30219 | | |
| VERSION | E30219.1 | GI:13017026 | |
| KEYWORDS | JP 1999318467-A/2. | | |
| SOURCE | unidentified | | |
| ORGANISM | unidentified | | |
| unclassified. | | | |
| REFERENCE | 1 | (bases 1 to 4402) | |
| AUTHORS | Sinichi, T. | | |
| TITLE | Shortened dystrophin | | |
| JOURNAL | Patent: JP 1999318467-A 2 24-NOV-1999; | | |
| | SCIENCE & TECH AGENCY, NATIONAL CENTER OF NEUROLOGY AND PSYCHIATRY | | |
| COMMENT | OS Unidentified | | |
| | PN JP 1999318467-A/2 | | |
| | PD 24-NOV-1999 | | |
| | PF 08-MAY-1998 JP 1998142134 | | |
| | PR | | |
| | PI SINICHI TAKEDA | | |
| | PC C12N15/09,A61K48/00,C12N15/00 | | |
| | CC Strandedness: Both; | | |
| | CC Topology: Linear; | | |
| | FH Key | Location/Qualifiers | |
| | FT source | 1..4402 | |
| | FT | /organism='Unidentified'. | |
| FEATURES | | Location/Qualifiers | |
| source | 1..4402 | /organism="unidentified" | |
| | | /mol_type="genomic DNA" | |
| | | /db_xref="taxon:32644" | |
| BASE COUNT | 1329 a | 1000 c | 1019 g |
| ORIGIN | | 1054 t | |
| Query Match | | 62.7%; | Score 1317; DB 6; Length 4402; |
| Best Local Similarity | | 81.6%; | Pred. No. 0; |
| Matches 1714; | Conservative | 0; | Mismatches 0; Indels 387; Gaps 1; |
| QY | 1 | GAGCTATGCCATACACACAGGCTGCTTATGTCACCACTCTGACCAAGTATTATCGTGGCCTT | 60 |
| Db | 1095 | GAGCTATGCCATACACACAGGCTGCTTATGTCACCACTCTGACCAAGTATTATCGTGGCCTT | 1154 |
| QY | 61 | TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCAATGATGGAGAG | 120 |
| Db | 1155 | TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCAATGATGGAGAG | 1214 |
| QY | 121 | TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC | 180 |
| Db | 1215 | TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC | 1274 |
| QY | 181 | TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA | 240 |
| Db | 1275 | TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA | 1334 |
| QY | 241 | CCAGTTTTCATACTCATGAGGGGTACATGATGCATTTGACAGCCCATCAGGGCCGGTTGG | 300 |
| Db | 1335 | CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG | 1394 |

| | | | |
|----|------|--|------|
| QY | 301 | TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTATCAGAAGATGAAGA | 360 |
| Db | 1395 | TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTATCAGAAGATGAAGA | 1454 |
| QY | 361 | AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC | 420 |
| Db | 1455 | AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC | 1514 |
| QY | 421 | TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT | 480 |
| Db | 1515 | TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT | 1574 |
| QY | 481 | GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAACAAAGAAAAATGGAGGAAGA | 540 |
| Db | 1575 | GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAACAAAGAAAAATGGAGGAAGA | 1634 |
| QY | 541 | GCCTCTTGGACCTGATCTTTGAAGACCTFAAAACGCCAAGTACAAACAATAAGGTGCTTCA | 600 |
| Db | 1635 | GCCTCTTGGACCTGATCTTTGAAGACCTFAAAACGCCAAGTACAAACAATAAGGTGCTTCA | 1694 |
| QY | 601 | AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT | 660 |
| Db | 1695 | AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT | 1754 |
| QY | 661 | TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG | 720 |
| Db | 1755 | TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTA----- | 1804 |
| QY | 721 | AGATCGATGGSCAAACATCTGTAGATGGACAGAAGACCGCTGGGTCTTTTACAAGACAG | 780 |
| Db | 1805 | ----- | 1804 |
| QY | 781 | TTCTGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCT | 840 |
| Db | 1805 | ----- | 1804 |
| QY | 841 | GAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAA | 900 |
| Db | 1805 | ----- | 1804 |
| QY | 901 | GCAGAACGATGTACATAGGGCCTTCAAGAGGGGAATTGAAAACCTAAAGAACCTGTAATCAT | 960 |
| Db | 1805 | ----- | 1804 |
| QY | 961 | GAGTACTCTTGAGACTGTACGAATATTCTTGACAGAGCAGCCCTTTGGAAGGACTAGAGAA | 1020 |
| Db | 1805 | ----- | 1804 |
| QY | 1021 | ACTCTACGAGGCCAGAGAGCTGCCTCTGTGAGGAGAGAGCCAGCAATGTCACTCGGCT | 1080 |
| Db | 1805 | ----- | 1804 |
| QY | 1081 | TCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTTGCACTCCGC | 1140 |
| Db | 1805 | -----AGGAGGTCAATACTAGTGGGAAAAATTGAACCTTGCACTCCGC | 1847 |
| QY | 1141 | TGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAAGACTCCAGGAACCTTCAAGAGGCCAC | 1200 |
| Db | 1848 | TGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAAGACTCCAGGAACCTTCAAGAGGCCAC | 1907 |
| QY | 1201 | GGATGAGCTGGACCTCAAGCTCGGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT | 1260 |
| Db | 1908 | GGATGAGCTGGACCTCAAGCTCGGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT | 1967 |
| QY | 1261 | GGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAGTCAAGGCACCTTCGAGG | 1320 |
| Db | 1968 | GGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAGTCAAGGCACCTTCGAGG | 2027 |
| QY | 1321 | AGAAATTGGCCTCTGAAAGAGAACCTGAGCCACGTCATGACCTTGCTCGCCAGCTTAC | 1380 |
| Db | 2028 | AGAAATTGGCCTCTGAAAGAGAACCTGAGCCACGTCATGACCTTGCTCGCCAGCTTAC | 2087 |
| QY | 1381 | CACCTTTGGGCATTGAGCTCTCACCCGTATTAACCTCAGCACCTCTGGAAGACCTGAACACCAG | 1440 |

Db 2088 CACCTTTGGGCAATTCAGCTCTCACCGTATACCTCAGCACTCTGGAAGACCTGAACACCAG 2147
Qy 1441 ATGGAAGCTTCTGCAGGTGGCGGTGAGGACCGAGTCAAGGAGCTGCATGAAGCCACAG 1500
Db 2148 ATGGAAGCTTCTGCAGGTGGCGGTGAGGACCGAGTCAAGGAGCTGCATGAAGCCACAG 2207
Qy 1501 GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCCCTGGGAGAG 1560
Db 2208 GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCCCTGGGAGAG 2267
Qy 1561 AGCCATCTCGCAAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACAACCTTGCTG 1620
Db 2268 AGCCATCTCGCAAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACAACCTTGCTG 2327
Qy 1621 GGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTTAGCTGACCTGAATAATGTCAGATT 1680
Db 2328 GGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTTAGCTGACCTGAATAATGTCAGATT 2387
Qy 1681 CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCT 1740
Db 2388 CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCT 2447
Qy 1741 CTTGAGCCTGTGACGTGCATGTGATGCTTGAACCTGGACGACACAACTCAAGCAAAATGACCA 1800
Db 2448 CTTGAGCCTGTGACGTGCATGTGATGCTTGAACCTGGACGACACAACTCAAGCAAAATGACCA 2507
Qy 1801 GCCCATGGATATCCTGCAGATTATTAAATTGTTTGACCACTATTATGACCGCCTGGAGCA 1860
Db 2508 GCCCATGGATATCCTGCAGATTATTAAATTGTTTGACCACTATTATGACCGCCTGGAGCA 2567
Qy 1861 AGAGCACAAACAATTGGTCAACGTCCTCTCTGCGTGGATATGTCTGAACTGGCTGCT 1920
Db 2568 AGAGCACAAACAATTGGTCAACGTCCTCTCTGCGTGGATATGTCTGAACTGGCTGCT 2627
Qy 1921 GAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGCTCTGCTTTTAAAACTGGCAT 1980
Db 2628 GAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGCTCTGCTTTTAAAACTGGCAT 2687
Qy 1981 CATTTCCCTGTGTAAGACACATTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGSC 2040
Db 2688 CATTTCCCTGTGTAAGACACATTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGSC 2747
Qy 2041 AAGTTCAACAGGATTTGTGACCGAGCGGCTGGCCCTCCTTCTGATGATTCTATCCA 2100
Db 2748 AAGTTCAACAGGATTTGTGACCGAGCGGCTGGCCCTCCTTCTGATGATTCTATCCA 2807
Qy 2101 A 2101
Db 2808 A 2808

RESULT 15
E30220
LOCUS E30220 4402 bp DNA linear PAT 18-JUN-2001
DEFINITION Shortened dystrophin.
ACCESSION E30220
VERSION E30220.1 GI:13017027
KEYWORDS JP 1999318467-A/3.
SOURCE unidentified
ORGANISM unidentified
REFERENCE unclassified.
AUTHORS 1 (bases 1 to 4402)
TITLE Sinichi,T.
JOURNAL Shortened dystrophin
COMMENT Patent: JP 1999318467-A 3 24-NOV-1999;
SCIENCE & TECH AGENCY,NATIONAL CENTER OF NEUROLOGY AND PSYCHIATRY
OS Unidentified
PN JP 1999318467-A/3
PD 24-NOV-1999
PF 08-MAY-1998 JP 1998142134
PR SINICHI TAKEDA

PC C12N15/09,A61K48/00,C12N15/00
CC Strandedness: Both;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..4402
FT /organism='Unidentified'.
FEATURES
source
1..4402
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 1339 a 1010 g 1069 t
ORIGIN
Query Match 57.0%; Score 1196.6; DB 6; Length 4402;
Best Local Similarity 78.3%; Pred. No. 1.3e-289;
Matches 1645; Conservative 0; Mismatches 69; Indels 387; Gaps 2;
Qy 1 GAGCTATGCTACACACAGGCTGCTTATGTCAACCCTCTGACCCCTACACGGAGCCCAATT 60
Db 1095 GAGCTATGCTACACACAGGCTGCTTATGTCAACCCTCTGACCCCTACACGGAGCCCAATT 1154
Qy 61 TCCTTCACAGCATTTGGAAGCTCCTGAAAGACAAGTCAATTTGGCAGTTTCATTGTGAGAG 120
Db 1155 TCCTTCACAGCATTTGGAAGCTCCTGAAAGACAAGTCAATTTGGCAGTTTCATTGTGAGAG 1214
Qy 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATTTATCGTGGCTTCTTTC 180
Db 1215 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATTTATCGTGGCTTCTTTC 1274
Qy 181 TGCTGAGGACACATTTGCAAGCACAAAGAGAGAGATTTCATATGATGTGAAAGTGGTGAAGA 240
Db 1275 TGCTGAGGACACATTTGCAAGCACAAAGAGAGAGATTTCATATGATGTGAAAGTGGTGAAGA 1334
Qy 241 CCAGTTTCATCTATCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 300
Db 1335 CCAGTTTCATCTATCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 1394
Qy 301 TAATATTTCTACAATTGGGAAGTAAGCTGATTTGGAACAGGAAATATATCAGAAGATGAAGA 360
Db 1395 TAATATTTCTACAATTGGGAAGTAAGCTGATTTGGAACAGGAAATATATCAGAAGATGAAGA 1454
Qy 361 AACTGAAGTACAAGACAGATGAATCTCTAAATTTCAAGATGGGAATGCCCTCAGGGTAGC 420
Db 1455 AACTGAAGTACAAGACAGATGAATCTCTAAATTTCAAGATGGGAATGCCCTCAGGGTAGC 1514
Qy 421 TAGCATGGAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 480
Db 1515 TAGCATGGAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 1574
Qy 481 GAAAGAGTTGAATGACTGGCTTAACAAAACAGAAAGAAACAGGAAATGGAGGAAGA 540
Db 1575 GAAAGAGTTGAATGACTGGCTTAACAAAACAGAAAGAAACAGGAAATGGAGGAAGA 1634
Qy 541 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACCGCCAAAGTACAACAACATAAGGTGCTTCA 600
Db 1635 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACCGCCAAAGTACAACAACATAAGGTGCTTCA 1694
Qy 601 AGAAGATCTAGAACAAAGCAAGTCAAGGTCAATTTCTCTCACTCAGATGGTGGTAGT 660
Db 1695 AGAAGATCTAGAACAAAGCAAGTCAAGGTCAATTTCTCTCACTCAGATGGTGGTAGT 1754
Qy 661 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 720
Db 1755 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG 1814
Qy 721 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAG 780
Db 1815 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 1874
Qy 781 TTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCT 840
Db 1875 CCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTATGTCATGGCTTTCAGA 1934

QY 841 GAAAGATGATGAATTAAAGCCGGCAGGCACCTATTGGAGCGGACTTTCCAGCAGTTCAGAA 900
Db ||||| |||||
1935 AAAAGAAGATGCAGT----- 1949
QY 901 GCAGAAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCAT 960
Db ----- 1949
QY 961 GAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAA 1020
Db ||||| |||||
1950 -----GAACAAGATTTCACACAACCTGGCTTTAAAGATCAAAATGAA 1989
QY 1021 ACTCTACAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCCAGAAATGTCACTCGGCT 1080
Db ||||| |||||
1990 ATGTTATCAAG----- 2000
QY 1081 TCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGGAATAATTGAACCTGCACCTCCGC 1140
Db ----- 2000
QY 1141 TGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCAC 1200
Db ----- 2000
QY 1201 GGATGAGCTGGACCTCAAGCTGGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT 1260
Db ----- 2000
QY 1261 GGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAGTCAAGGCACCTTCGAGG 1320
Db ||||| |||||
2001 -----TCTCGAGAAAGTCAAGGCACCTTCGAGG 2027
QY 1321 AGAAATTGCGCCTCTGAAAGAGAAACGTGAGGCCACGTCATGACCTTGCTCGCCAGCTTAC 1380
Db ||||| |||||
2028 AGAAATTGCGCCTCTGAAAGAGAAACGTGAGGCCACGTCATGACCTTGCTCGCCAGCTTAC 2087
QY 1381 CACTTTGGGCATTGAGCTCTCACCGTATAACTCAGCACTCTGGAAGACCTGAACACCAG 1440
Db ||||| |||||
2088 CACTTTGGGCATTGAGCTCTCACCGTATAACTCAGCACTCTGGAAGACCTGAACACCAG 2147
QY 1441 ATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAAGGAGCTGCATGAAGCCACACAG 1500
Db ||||| |||||
2148 ATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAAGGAGCTGCATGAAGCCACACAG 2207
QY 1501 GGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCGTGCCAGGGTCCCTGGGAGAG 1560
Db ||||| |||||
2208 GGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCGTGCCAGGGTCCCTGGGAGAG 2267
QY 1561 AGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACAACCTTGCTG 1620
Db ||||| |||||
2268 AGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACAACCTTGCTG 2327
QY 1621 GGACCATCCCAAAATGACAGAGCTTACCAGTCTTTAGCTGACCTGAAATAATGTCAGATT 1680
Db ||||| |||||
2328 GGACCATCCCAAAATGACAGAGCTTACCAGTCTTTAGCTGACCTGAAATAATGTCAGATT 2387
QY 1681 CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAGGCCCTTTGCTTGGATCT 1740
Db ||||| |||||
2388 CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAGGCCCTTTGCTTGGATCT 2447
QY 1741 CTTGAGCCTGTGAGCTGCATGTGATGCCCTTGGACCAGCAACCTCAAGCAAAATGACCA 1800
Db ||||| |||||
2448 CTTGAGCCTGTGAGCTGCATGTGATGCCCTTGGACCAGCAACCTCAAGCAAAATGACCA 2507
QY 1801 GCCCATGGATATCCTGCAGATTATTAAATTGTTTGACCACCTATTATGACCGCCTGGAGCA 1860
Db ||||| |||||
2508 GCCCATGGATATCCTGCAGATTATTAAATTGTTTGACCACCTATTATGACCGCCTGGAGCA 2567
QY 1861 AGAGCACAACAAATTGGTCAACGTCCTCTCTCGTGGATATGTCTGAACCTGGCTGCT 1920
Db ||||| |||||
2568 AGAGCACAACAAATTGGTCAACGTCCTCTCTCGTGGATATGTCTGAACCTGGCTGCT 2627

QY 1921 GAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGTCTCTCTTTTAAAACTGGCAT 1980
Db ||||| |||||
2628 GAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGTCTCTCTTTTAAAACTGGCAT 2687
QY 1981 CATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGC 2040
Db ||||| |||||
2688 CATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGC 2747
QY 2041 AAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCCTCCTTCTGCAATGATTCCTATCCA 2100
Db ||||| |||||
2748 AAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCCTCCTTCTGCAATGATTCCTATCCA 2807
QY 2101 A 2101
Db 2808 A 2808

Search completed: February 1, 2004, 23:51:21
Job time : 7497.6 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

3M nucleic - nucleic search, using sw model

Run on: February 1, 2004, 10:13:25 ; Search time 499.045 Seconds
(without alignments)
11364.749 Million cell updates/sec

Title: US-09-845-416-12_COPY_900_3000
Perfect score: 2101
Sequence: 1 gagctatgcctacacacagg.....ttctgcatgtattctatccaa 2101

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query % | | Length | DB ID | Description |
|------------|---------|-------|--------|-------|--------------------|
| | Score | Match | | | |
| 1 | 2101 | 100.0 | 3510 | 24 | Human dystrophin m |
| 2 | 2101 | 100.0 | 4476 | 24 | Adeno-associated v |
| 3 | 1760.4 | 83.8 | 5339 | 24 | DNA encoding mini- |
| 4 | 1750.8 | 83.3 | 3531 | 24 | Human dystrophin m |
| 5 | 1750.8 | 83.3 | 4498 | 24 | Adeno-associated v |
| 6 | 1743 | 83.0 | 3858 | 24 | Human dystrophin m |
| 7 | 1743 | 83.0 | 4825 | 24 | Adeno-associated v |
| 8 | 1743 | 83.0 | 4848 | 24 | Adeno-associated v |

| | | | | | | |
|----|--------|------|-------|----|----------|--------------------|
| 9 | 1743 | 83.0 | 5060 | 24 | AAD37264 | Adeno-associated v |
| 10 | 1641 | 78.1 | 4414 | 24 | AAD37260 | Adeno-associated v |
| 11 | 1629 | 77.5 | 3446 | 24 | AAD37242 | Human dystrophin m |
| 12 | 1629 | 77.5 | 5462 | 24 | ABK81999 | DNA encoding mini- |
| 13 | 1611.6 | 76.7 | 5417 | 24 | ABK81997 | DNA encoding mini- |
| 14 | 1602 | 75.2 | 3999 | 24 | AAD37234 | Human dystrophin m |
| 15 | 1602 | 76.2 | 4966 | 24 | AAD37256 | Adeno-associated v |
| 16 | 1602 | 76.2 | 4990 | 24 | AAD37262 | Adeno-associated v |
| 17 | 1419 | 67.5 | 4182 | 24 | AAD37230 | Human dystrophin m |
| 18 | 1419 | 67.5 | 5149 | 24 | AAD37255 | Adeno-associated v |
| 19 | 1328.8 | 63.2 | 2169 | 24 | AAD37232 | Human dystrophin r |
| 20 | 1328.8 | 63.2 | 5952 | 22 | AAD06794 | Human dystrophin g |
| 21 | 1328.8 | 63.2 | 8689 | 24 | ABK82000 | DNA encoding mini- |
| 22 | 1328.8 | 63.2 | 11058 | 24 | AAD37229 | Human dystrophin p |
| 23 | 1328.8 | 63.2 | 11241 | 24 | ABK82005 | CDNA encoding huma |
| 24 | 1328.8 | 63.2 | 11443 | 24 | ABK82002 | DNA encoding mini- |
| 25 | 1328.8 | 63.2 | 12923 | 10 | AA90338 | Sequence of human |
| 26 | 1328.8 | 63.2 | 13957 | 24 | ABT10904 | Human breast cance |
| 27 | 1328.8 | 63.2 | 13957 | 24 | ABK69900 | Human dystrophin g |
| 28 | 1328.8 | 63.2 | 13957 | 24 | ABN95786 | Gene #2284 used to |
| 29 | 1328.8 | 63.2 | 13957 | 24 | ABK81959 | CDNA encoding huma |
| 30 | 1324 | 63.0 | 1821 | 24 | AAD37241 | Human dystrophin r |
| 31 | 1317.8 | 62.7 | 13977 | 24 | ABK70403 | Human bone remodel |
| 32 | 1317 | 62.7 | 4402 | 21 | AZ48567 | A rod shortened dy |
| 33 | 1196.6 | 57.0 | 4402 | 21 | AZ48568 | A rod shortened dy |
| 34 | 1149.6 | 54.7 | 13815 | 24 | ABK81960 | CDNA encoding mous |
| 35 | 1149.6 | 54.7 | 13815 | 24 | AB199799 | Mouse ischaemic co |
| 36 | 1149.6 | 54.7 | 19307 | 17 | AAT27558 | Shuttle vector pAd |
| 37 | 1148 | 54.6 | 13815 | 19 | AAV18885 | Mus musculus dystr |
| 38 | 937 | 44.6 | 1434 | 24 | AAD37243 | Human dystrophin r |
| 39 | 785.4 | 37.4 | 1991 | 24 | AAD37231 | Human dystrophin N |
| 40 | 777 | 37.0 | 1667 | 24 | AAD37235 | Human dystrophin N |
| 41 | 718.8 | 34.2 | 4075 | 21 | AZ48569 | A rod shortened dy |
| 42 | 667.2 | 31.8 | 3747 | 21 | AZ48566 | A rod shortened dy |
| 43 | 618.6 | 29.4 | 3275 | 10 | AA97129 | Partial sequence o |
| 44 | 613.6 | 29.2 | 3163 | 21 | AZ48571 | A rod shortened dy |
| 45 | 554.2 | 26.4 | 11096 | 24 | ABK81962 | CDNA encoding mous |

ALIGNMENTS

RESULT 1

AAD37240
ID AAD37240 standard; DNA; 3510 BP.

XX AAD37240;

XX 21-AUG-2002 (first entry)

DE Human dystrophin minigene delta3510.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.

XX Homo sapiens.

XX WO200183695-A2.

XX 08-NOV-2001.

PF 27-APR-2001; 2001WO-US13677.

XX 28-APR-2000; 2000US-200777P.

XX (XIAO/) XIAO X.

XX XIAO X;

XX WPI; 2002-049342/06.

PT New dystrophin minigene for treating Duchenne or Becker muscular

QY 961 GAGTACTCTTGAGACTGTACCAATATTTCTGACAGACAGCCCTTTGGAGGAGCTAGAGAA 1020
Db 2616 GAGTACTCTTGAGACTGTACCAATATTTCTGACAGACAGCCCTTTGGAGGAGCTAGAGAA 2675
QY 1021 ACTCTACAGGAGCCCGAGAGCTGCCTCCTGAGGAGAGCCCGAAGTGTCACTCGGCT 1080
Db 2676 ACTCTACAGGAGCCCGAGAGAGCTGCCTCCTGAGGAGAGAGCCCGAAGTGTCACTCGGCT 2735
QY 1081 TCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTGCCTCCGC 1140
Db 2736 TCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTGCCTCCGC 2795
QY 1141 TGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCAC 1200
Db 2796 TGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCAC 2855
QY 1201 GGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT 1260
Db 2856 GGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT 2915
QY 1261 GGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGG 1320
Db 2916 GGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGG 2975
QY 1321 AGAAATTGGCCCTCTGAAAGAGAAAGCTGAGCCACGTCATGACCTTGCTCGCCAGCTTAC 1380
Db 2976 AGAAATTGGCCCTCTGAAAGAGAAAGCTGAGCCACGTCATGACCTTGCTCGCCAGCTTAC 3035
QY 1381 CACTTTGGSCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAG 1440
Db 3036 CACTTTGGSCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAG 3095
QY 1441 ATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGAGCTGCATGAAGCCACAG 1500
Db 3096 ATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGAGCTGCATGAAGCCACAG 3155
QY 1501 GGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGCTCTGCCAGGTCCTGGGAGAG 1560
Db 3156 GGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGCTCTGCCAGGTCCTGGGAGAG 3215
QY 1561 AGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACCACGAGACTCAAAACAACTTGCTG 1620
Db 3216 AGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACCACGAGACTCAAAACAACTTGCTG 3275
QY 1621 GGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAATGTCAGATT 1680
Db 3276 GGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAATGTCAGATT 3335
QY 1681 CTCAGCTTATAGACTGCCATGAAACTCCGAAGACTGCAGAGCCCTTTGCTTGGATCT 1740
Db 3336 CTCAGCTTATAGACTGCCATGAAACTCCGAAGACTGCAGAGCCCTTTGCTTGGATCT 3395
QY 1741 CTTGAGCCTGTACAGCTGCATGTATGCTTGGACCAGCACAACTCAAGCAAAATGACCA 1800
Db 3396 CTTGAGCCTGTACAGCTGCATGTATGCTTGGACCAGCACAACTCAAGCAAAATGACCA 3455
QY 1801 GCCCATGGATATCCTGCAGATTATTAATTGTTGACCACTATTTATGACCGCCTGGAGCA 1860
Db 3456 GCCCATGGATATCCTGCAGATTATTAATTGTTGACCACTATTTATGACCGCCTGGAGCA 3515
QY 1861 AGAGCACAACAAATTTGGTCAACGTCCTCTGCTGGATATGTCTGAACTGGCTGCT 1920
Db 3516 AGAGCACAACAAATTTGGTCAACGTCCTCTCTGCTGGATATGTCTGAACTGGCTGCT 3575
QY 1921 GAATGTTTATGATACGGGACGAACAGGAGGATCCGTGCTGCTTTTAAAACTGGCAT 1980
Db 3576 GAATGTTTATGATACGGGACGAACAGGAGGATCCGTGCTGCTTTTAAAACTGGCAT 3635
QY 1981 CATTTCCCTGTGTAAGCACATTTTGGAAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGC 2040
Db 3636 CATTTCCCTGTGTAAGCACATTTTGGAAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGC 3695
QY 2041 AAGTTCAACAGGATTTTGTGACCAGCGGCTGGGCCCTCCTTCTGCGATGATTCTATCCA 2100

Db 3696 AAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCCTCCTTCTGCGATGATTCTATCCA 3755
QY 2101 A 2101
Db 3756 A 3756
RESULT 3
ABK81998
ID ABK81998 standard; DNA; 5339 BP.
XX
AC ABK81998;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding mini-dystrophin protein deltaR2-R21.
XX
KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200229056-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US31126.
XX
PR 06-OCT-2000; 2000US-238848P.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Chamberlain JS, Harper SQ;
XX
DR WPI; 2002-435334/46.
XX
PT A composition for preparing therapeutic drugs, has a mini-dystrophin peptide comprising a specific number of spectrin-like repeat domains, or a nucleic acid sequence encoding the mini-dystrophin peptide -
XX
PS Example 6; Fig 13; 145pp; English.
XX
CC The invention describes a composition comprising a mini-dystrophin peptide comprising a spectrin-like repeat domain, where the domain comprises n spectrin-like repeats, and contains no more than n spectrin-like repeats, where n is an even number between 4-24, or a nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the polynucleotide encoding it is useful as a medicament, for preparing a drug for therapeutic application and in the preparation of a composition for treatment of muscle disease, e.g. Duchenne's muscular dystrophy (DMD). This sequence represents a mini-dystrophin sequence of the invention.
XX
SQ Sequence 5339 BP; 1638 A; 1191 C; 1187 G; 1323 T; 0 other;

Query Match 83.8%; Score 1760.4; DB 24; Length 5339;
Best Local Similarity 90.3%; Pred.No. 0;
Matches 1910; Conservative 0; Mismatches 191; Indels 15; Gaps 2;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCCTCTGACCTACACGGAGCCCAATT 60
Db 1099 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCCTCTGACCTACACGGAGCCCAATT 1158
QY 61 TCCTTCACAGCAATTGGAAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 120
Db 1159 TCCTTCACAGCAATTGGAAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGTATTATCAAAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTC 180
Db 1219 TGAAGTAAACCTGGACCGTATTATCAAAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTC 1278

2Y 181 TGCTGAGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTCGTGAAAGA 240
Db 1279 TGCTGAGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTCGTGAAAGA 1338
2Y 241 CCAGTTTCATACATGAGGGGTACATGATGGATTGACAGCCCCATCAGGGCCGGTTGG 300
Db 1339 CCAGTTTCATACATGAGGGGTACATGATGGATTGACAGCCCCATCAGGGCCGGTTGG 1398
2Y 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAAGATGAAGA 360
Db 1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAAGATGAAGA 1458
2Y 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCCTCAGGGTAGC 1518
2Y 421 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAAACT 480
Db 1519 TAGCATGGAAAAACAAAGCAATTACATCATAGATTACTGCAACAGTTCCCCCTGGACCT 1578
2Y 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGGAAAAATGGAGGAAGA 540
Db 1579 GGAAGAGTTTCTTGCCCTGGCTTACAGAAGCTGAAACAACACTGCCAATGTCTCAGAGATGC 1638
2Y 541 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAACAACATAAAGTCTTCA 600
Db 1639 TACCCGTAAAGAAAGGCTCCTAGAAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATG 1698
2Y 601 AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 660
Db 1699 GCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAAA 1758
2Y 661 TGATGA-----ATCTAGTGGAGATCAAGCAACTGCTGCTTTTGGAGAACAACT 708
Db 1759 CAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTCAACAATAGGTC 1818
2Y 709 TAAGG---TATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGT 765
Db 1819 TTTGGATAACATGAACCTTCAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACAATAGGTC 1878
2Y 766 TCTTTTACAAGACAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCT 825
Db 1879 CCATTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCT 1938
2Y 826 GGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGCGCACTT 885
Db 1939 GGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGCGCACTT 1998
2Y 886 TCCAGCAGTTCAAGACAGAACCGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAA 945
Db 1999 TCCAGCAGTTCAAGACAGAACCGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAA 2058
2Y 946 AGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTT 1005
Db 2059 AGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTT 2118
2Y 1006 GGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCA 1065
Db 2119 GGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCA 2178
2Y 1066 GAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATT 1125
Db 2179 GAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATT 2238
2Y 1126 GAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGA 1185
Db 2239 GAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGA 2298
2Y 1186 ACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGG 1245
Db 2299 ACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGG 2358
2Y 1246 ATCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAGATCACCTCGAGAAAGT 1305

Db 2359 ATCTTGCGAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAAAGT 2418
2Y 1306 CAAGGCACCTTCGAGGAGAAATTCGCCTCTGAAAGAGAAAGTGAAGCCACGTCAATGACCT 1365
Db 2419 CAAGGCACCTTCGAGGAGAAATTCGCCTCTGAAAGAGAAAGTGAAGCCACGTCAATGACCT 2478
2Y 1366 TGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACCTCTGGA 1425
Db 2479 TGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACCTCTGGA 2538
2Y 1426 AGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCCTCGAGGACCGAGTCAGGCAGCT 1485
Db 2539 AGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCCTCGAGGACCGAGTCAGGCAGCT 2598
2Y 1486 GCATGAAGCCCAACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTCTGTCCA 1545
Db 2599 GCATGAAGCCCAACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTCTGTCCA 2658
2Y 1546 GGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCACGAGAC 1605
Db 2659 GGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCACGAGAC 2718
2Y 1606 TCAACAAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGTGAACCT 1665
Db 2719 TCAACAAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGTGAACCT 2778
2Y 1666 GAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAGGC 1725
Db 2779 GAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAGGC 2838
2Y 1726 CTTTGTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCTTGGACCAGCAACCT 1785
Db 2839 CTTTGTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCTTGGACCAGCAACCT 2898
2Y 1786 CAAGCAAAATGACCGCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTA 1845
Db 2899 CAAGCAAAATGACCGCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTA 2958
2Y 1846 TGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCCTCTCTCGTGGATATGTG 1905
Db 2959 TGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCCTCTCTCGTGGATATGTG 3018
2Y 1906 TCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCTGTCTC 1965
Db 3019 TCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCTGTCTC 3078
2Y 1966 TTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGATACCT 2025
Db 3079 TTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGATACCT 3138
2Y 2026 TTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGGCTCTCTTCT 2085
Db 3139 TTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGGCTCTCTTCT 3198
2Y 2086 GCATGATTCTATCCAA 2101
Db 3199 GCATGATTCTATCCAA 3214

RESULT 4

AAD37238
ID AAD37238 standard; DNA; 3531 BP.

XX AAD37238;

XX

DT 21-AUG-2002 (first entry)

XX

DE Human dystrophin minigene delta3531.

XX

KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

KW Becker muscular dystrophy; ds.

XX OS Homo sapiens.
XX PN WO200183695-A2.
XX PD 08-NOV-2001.
XX 27-APR-2001; 2001WO-US13677.
XX PF 28-APR-2000; 2000US-200777P.
XX PR (XIAO/) XIAO X.
XX PA XIAO X;
XX PI WPI; 2002-049342/06.
XX DR New dystrophin minigene for treating Duchenne or Becker muscular
XX PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
XX PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
XX PT dystrophin gene -
XX PS Example 1; Page 50-51; 71pp; English.
XX CC The present invention relates to an isolated nucleotide sequence encoding
XX CC a dystrophin minigene. The minigene comprises N-terminal or modified
XX CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
XX CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
XX CC invention also relates to a recombinant adeno-associated virus (AAV)
XX CC comprising dystrophin minigene operably linked to an expression control
XX CC element. The dystrophin minigene in operable linkage with an expression
XX CC control element, in a recombinant adeno-associated virus or retrovirus is
XX CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
XX CC dystrophy (BMD) in a mammalian subject. The present sequence is human
XX CC dystrophin minigene delta3531 containing nucleotides 1-1341 (N-terminus,
XX CC hinge H1 and rod R1), 8059-10227 (rods R22, R23 and R24, hinge H4 and
XX CC CR domain) and 11047-11058 (dystrophin last 3 amino acids).
XX SQ Sequence 3531 BP; 1071 A; 809 C; 824 G; 827 T; 0 other;
Query Match 83.3%; Score 1750.8; DB 24; Length 3531;
Best Local Similarity 90.2%; Pred. No. 0;
Matches 1914; Conservative 0; Mismatches 187; Indels 21; Gaps 3;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACTCTGACCCCTACACGGAGCCCAT 60
DB 900 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACTCTGACCCCTACACGGAGCCCAT 959
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 120
DB 960 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 1019
QY 121 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
DB 1020 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1079
QY 181 TGCTGAGGACACATTTCAAGCACACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
DB 1080 TGCTGAGGACACATTTCAAGCACACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1139
QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
DB 1140 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1199
QY 301 TAATATTCTACAAATTGGGAAGTAAGCTGATTGGACAGGAAAATTATCAGAAGATGAAGA 360
DB 1200 TAATATTCTACAAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTATCAGAAGATGAAGA 1259
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCTCAGGGTAGC 420
DB 1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCTCAGGGTAGC 1319
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTT-----TTAATGGATCTCCAGAAATCA 474

DB 1320 TAGCATGGAAAAACAAAGCAATTTACATAGAACTCATAGATTACTGCAACAGTTCCCCCT 1379
QY 475 GAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAAGAAATGGA 534
DB 1380 GGACCTGGAAAAAGTTTCTTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACA 1439
QY 535 GGAAGAGCCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACAATAAGGT 594
DB 1440 GGATGCTACCCGTAAAGAAAGCTCCTAGAAGACTCCAAAGGGAGTAAAGAGCTGATGAA 1499
QY 595 GCTTCAAGAAGATCTAGAACAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGGTGT 654
DB 1500 ACAATGGCAAGACCTCCAAAGGTGAAATTGAAGCTCACACAGATGTTTTATCACAACCTGGA 1559
QY 655 GGTAG-----TTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGGAAAGA 702
DB 1560 TGAAAAACAGCCAAAAATCCTGAGATCCCTGGAGGTTCCGATGATGCGAGTCTCTGTACA 1619
QY 703 ACAACTTAAGG---TATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGACCG 759
DB 1620 AAGACGTTTGGATAACATGAACATTCAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACAT 1679
QY 760 CTGGGTTCTTTTACAAGACAGTTCTGACCAGTGGAGCGTCTGCACCTTTCTCTGCAGGA 819
DB 1680 TAGTCCCATTTGGAAGCCAGTCTGACCAGTGGAGCGTCTGCACCTTTCTCTGCAGGA 1739
QY 820 ACTTCTGCTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGG 879
DB 1740 ACTTCTGCTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGG 1799
QY 880 CGACTTTCAGCAGTTTCAAGACAGAAACGATGTACATAGGGCTTCAAGAGGGAATTGAA 939
DB 1800 CGACTTTCAGCAGTTTCAAGACAGAAACGATGTACATAGGGCTTCAAGAGGGAATTGAA 1859
QY 940 AACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 999
DB 1860 AACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 1919
QY 1000 GCCTTTTGAAGGACTAGAGAACTCTACAGGAGCCCAAGAGCTGCCTCTCTGAGGAGAG 1059
DB 1920 GCCTTTTGAAGGACTAGAGAACTCTACAGGAGCCCAAGAGCTGCCTCTCTGAGGAGAG 1979
QY 1060 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 1119
DB 1980 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 2039
QY 1120 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACT 1179
DB 2040 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACT 2099
QY 1180 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT 1239
DB 2100 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT 2159
QY 1240 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAAGATCACCTCGA 1299
DB 2160 CAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAAGATCACCTCGA 2219
QY 1300 GAAAGTCAAGGCACCTTCGAGGAGAAATTGGCGCTCTGAAAAGAGAACGTGAGCCACGTCAA 1359
DB 2220 GAAAGTCAAGGCACCTTCGAGGAGAAATTGGCGCTCTGAAAAGAGAACGTGAGCCACGTCAA 2279
QY 1360 TGACCTTGTCTGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAC 1419
DB 2280 TGACCTTGTCTGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAC 2339
QY 1420 TCTGGAAGACCTGAAACACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAG 1479
DB 2340 TCTGGAAGACCTGAAACACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAG 2399
QY 1480 GCAGCTGCATGAAGCCCAACAGGGACTTTGGTTCAGCATCTTCAGCACCTTTCTTCCACGTC 1539

Db 2400 GCAGCTGCATGAAGCCACACAGGACTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTC 2459
QY 1540 TGTCCAGGGTCCCTGGGAGAGAGCCACTCTGCCAAACAAAGTGCCTACTATATCAACCA 1599
Db 2460 TGTCCAGGGTCCCTGGGAGAGAGCCACTCTGCCAAACAAAGTGCCTACTATATCAACCA 2519
QY 1600 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGTCTACCAGTCTTTAGC 1659
Db 2520 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGTCTACCAGTCTTTAGC 2579
QY 1660 TGACCTGAATAATGTGCAGATTCTCAGCTTATAGGACTGCCATGAAGTCCGAAGACTGCA 1719
Db 2580 TGACCTGAATAATGTGCAGATTCTCAGCTTATAGGACTGCCATGAAGTCCGAAGACTGCA 2639
QY 1720 GAAGGCCCTTTGCTGGATCTCTTGAGCCTGTGCAGCTGCATGTGATGCTTTGGACCAGCA 1779
Db 2640 GAAGGCCCTTTGCTGGATCTCTTGAGCCTGTGCAGCTGCATGTGATGCTTTGGACCAGCA 2699
QY 1780 CAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCGAGATTATTAATTGTTTGACCAC 1839
Db 2700 CAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCGAGATTATTAATTGTTTGACCAC 2759
QY 1840 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGGA 1899
Db 2760 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGGA 2819
QY 1900 TATGTGCTGAATGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1959
Db 2820 TATGTGCTGAATGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 2879
QY 1960 CCTGTCTTTTAAACTGGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAG 2019
Db 2880 CCTGTCTTTTAAACTGGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAG 2939
QY 2020 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCGCAGGCTGGGCCT 2079
Db 2940 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCGCAGGCTGGGCCT 2999
QY 2080 CCTTCTGCATGATCTATCCAA 2101
Db 3000 CCTTCTGCATGATCTATCCAA 3021

RESULT 5
AAD37258
ID AAD37258 standard; DNA; 4498 BP.
XX AAD37258;
XX
XX
DT 21-AUG-2002 (first entry)
XX Adeno-associated virus vector plasmid, AAV-MCK-3531.
DE Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX Chimeric - Homo sapiens.
OS Chimeric - Unidentified.
XX WO200183695-A2.
XX
XX 08-NOV-2001.
XX
XX 27-APR-2001; 2001WO-US13677.
XX
XX 28-APR-2000; 2000US-200777P.
XX (XIAO/) XIAO X.
XX XIAO X;
PI WPI; 2002-049342/06.
XX DR

XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 62-63; 71pp; English.
XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence.
XX
SQ Sequence 4498 BP; 1251 A; 1118 C; 1123 G; 1006 T; 0 other;
Query Match 83.3%; Score 1750.8; DB 24; Length 4498;
Best Local Similarity 90.2%; Pred. No. 0;
Matches 1914; Conservative 0; Mismatches 187; Indels 21; Gaps 3;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCTCTGACCCCTACCGAGCCCAT 60
Db 1657 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCTCTGACCCCTACCGAGCCCAT 1716
QY 61 TCCTTACAGCAATTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 120
Db 1717 TCCTTACAGCAATTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 1776
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTCGTGGCTTCTTTC 180
Db 1777 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTCGTGGCTTCTTTC 1836
QY 181 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGGAAAGTGGTGAAGA 240
Db 1837 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGGAAAGTGGTGAAGA 1896
QY 241 CCAGTTTTCATCTATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTGG 300
Db 1897 CCAGTTTTCATCTATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTGG 1956
QY 301 TAATATTCTACAATTGGGAAGTAAAGCTGATGGAAACAGGAAATTTATCAGAAGATGAAGA 360
Db 1957 TAATATTCTACAATTGGGAAGTAAAGCTGATGGAAACAGGAAATTTATCAGAAGATGAAGA 2016
QY 361 AACTGAAGTACAAGACAGATGAATCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
Db 2017 AACTGAAGTACAAGACAGATGAATCTCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 2076
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTT-----TTAATGGATCTCCAGAATCA 474
Db 2077 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTACTGCAACAGTTCCCCCT 2136
QY 475 GAAACTGAAAGAGTTGAATGACTGGCTACAAAAACAGAAGAAAGCAAGGAAATGGA 534
Db 2137 GGACCTGGAAAAAGTTTCTTGGCTTACAGAAGCTGAAAACAACCTGCCAATGTCTCTACA 2196
QY 535 GGAAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACACAACATAAGGT 594
Db 2197 GGATGCTACCCGTAAGGAAAGGCTCCTAGAGACTCCAAGGGAGTAAAGAGCTGATGAA 2256
QY 595 GCTTCAAGAAGATCTAGAACAAGAACAAAGTACAGGCTCAATTTCTCTACTACATGGTGGT 654
Db 2257 ACAATGGCAAGACCTCCAAGGTGAAATTTGAAGTCAACAGATGTTTATCACAACCTGGA 2316
QY 655 GGTAG-----TTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGA 702

Db 2317 TGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAAGGTTCCGATGATGACGTCCTGTTACA 2376
QY 703 ACAACTTAAGS--TATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCG 759
Db 2377 AAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACITTCGGAAAAAGTCTCTCAACAT 2436
QY 760 CTGGGTTCTTTTACAAGACAGATTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGACAGGA 819
Db 2437 TAGGTCCCATTTGGAAGCCAGTTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGACAGGA 2496
QY 820 ACTTCTGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGG 879
Db 2497 ACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGG 2556
QY 880 CGACTTTCAGCAGATTGAGAAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGGAATTGAA 939
Db 2557 CGACTTTCAGCAGATTGAGAAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGGAATTGAA 2616
QY 940 AACTAAAGAACCTGTGAATCATGAGTACTCTTGTAGACTGTACGAAATATTTCTGACAGAGCA 999
Db 2617 AACTAAAGAACCTGTGAATCATGAGTACTCTTGTAGACTGTACGAAATATTTCTGACAGAGCA 2676
QY 1000 GCCTTTGGAAGGACTAGAGAAACTCTACAGAGGCCAGAGAGTGCCTCCTGAGGAGAG 1059
Db 2677 GCCTTTGGAAGGACTAGAGAAACTCTACAGAGGCCAGAGAGTGCCTCCTGAGGAGAG 2736
QY 1060 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGA 1119
Db 2737 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGA 2796
QY 1120 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 1179
Db 2797 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 2856
QY 1180 CCAGGAACCTCAAGAGGCCACGGATGAGCTGACCTCAAGCTGCGCAAGCTGAGGTGAT 1239
Db 2857 CCAGGAACCTCAAGAGGCCACGGATGAGCTGACCTCAAGCTGCGCAAGCTGAGGTGAT 2916
QY 1240 CAAGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCAAGATCACCTCGA 1299
Db 2917 CAAGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCAAGATCACCTCGA 2976
QY 1300 GAAAGTCAAGGCACCTTCAGAGAGAAATTGCGCTCTGAAAGAGAACGTGAGCCACGTCAA 1359
Db 2977 GAAAGTCAAGGCACCTTCAGAGAGAAATTGCGCTCTGAAAGAGAACGTGAGCCACGTCAA 3036
QY 1360 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAACCTCAGCAC 1419
Db 3037 TGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAACCTCAGCAC 3096
QY 1420 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAG 1479
Db 3097 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAG 3156
QY 1480 GCAGCTGCATGAAGCCCAAGGGGACTTTGGTCCAGCATCTCAGCACCTTTCTTCCACGTC 1539
Db 3157 GCAGCTGCATGAAGCCCAAGGGGACTTTGGTCCAGCATCTCAGCACCTTTCTTCCACGTC 3216
QY 1540 TGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 1599
Db 3217 TGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 3276
QY 1600 CGAGCTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 1659
Db 3277 CGAGCTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 3336
QY 1660 TGACCTGAATAATGTGATTTCTGAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 1719
Db 3337 TGACCTGAATAATGTGATTTCTGAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3396
QY 1720 GAAGGCCCTTGTGTTGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCCTTGGACCAAGCA 1779
Db 3397 GAAGGCCCTTGTGTTGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCCTTGGACCAAGCA 3456

QY 1780 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC 1839
Db 3457 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC 3516
QY 1840 TATTATGACCGCCTGGAGCAAGAGCACAAACAATTTGGTCAACGTCCCTCTCTGCGTGA 1899
Db 3517 TATTATGACCGCCTGGAGCAAGAGCACAAACAATTTGGTCAACGTCCCTCTCTGCGTGA 3576
QY 1900 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACCGGACGAACAGGAGGATCCGTGT 1959
Db 3577 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACCGGACGAACAGGAGGATCCGTGT 3636
QY 1960 CCTGTCTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 2019
Db 3637 CCTGTCTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 3696
QY 2020 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGSCCT 2079
Db 3697 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGSCCT 3756
QY 2080 CCTTCTGCATGATTCTATCCAA 2101
Db 3757 CCTTCTGCATGATTCTATCCAA 3778

RESULT 6

AAD37237
ID AAD37237 standard; DNA; 3858 BP.
XX
AC AAD37237;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta3849.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 48-49; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human

CC dystrophin minigene delta3849 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1, R2), 8059-10227 (rods R22, R23 and R24, hinge H4
CC and CR domain) and 11047-11058 (dystrophin last 3 amino acids).
XX
SQ Sequence 3858 BP; 1189 A; 866 C; 905 G; 898 T; 0 other;

Query Match 83.0%; Score 1743; DB 24; Length 3858;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 348; Gaps 1;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTACACACCTCTGACCTACACGGAGCCCAT 60
Db |||||
QY 900 GAGCTATGCCTACACACAGGCTGCTTATGTACACACCTCTGACCTACACGGAGCCCAT 959
Db |||||
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGCAAGTCAATTTGGCAGTTCATTTGATGGAGAG 120
Db |||||
QY 960 TCCTTCACAGCATTTGGAAGCTCCTGAAGCAAGTCAATTTGGCAGTTCATTTGATGGAGAG 1019
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTTATCGTGGCTTCTTTC 180
Db |||||
QY 1020 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTTATCGTGGCTTCTTTC 1079
QY 181 TGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db |||||
QY 1080 TGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1139
QY 241 CCAGTTTCATACCTATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
Db |||||
QY 1140 CCAGTTTCATACCTATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1199
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATGGAAACAGGAAATTTATCAGAAGATGAAGA 360
Db |||||
QY 1200 TAATATTCTACAATTGGGAAGTAAGCTGATGGAAACAGGAAATTTATCAGAAGATGAAGA 1259
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
Db |||||
QY 1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1319
QY 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480
Db |||||
QY 1320 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 1379
QY 481 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAGAAAGAAACAAGGAAATGGAGGAAGA 540
Db |||||
QY 1380 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAGAAAGAAACAAGGAAATGGAGGAAGA 1439
QY 541 GCCTCTTGACCTGATCTTGAAGACCTAAAACGCAAGTACAAACAACATAAGGTGCTTCA 600
Db |||||
QY 1440 GCCTCTTGACCTGATCTTGAAGACCTAAAACGCAAGTACAAACAACATAAGGTGCTTCA 1499
QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTAGT 660
Db |||||
QY 1500 AGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTAGT 1559
QY 661 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTGGG 720
Db |||||
QY 1560 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTGGG 1619
QY 721 AGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGGGTTCTTTTACAAGA --- 777
Db |||||
QY 1620 AGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGGGTTCTTTTACAAGACAC 1679
QY 778 ----- 777
Db 1680 TCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAAGTTTCTTGGCTGGCTTACAGA 1739
QY 778 ----- 777
Db 1740 AGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTGAAGAAAGGCTCCTAGAAGA 1799
QY 778 ----- 777
Db 1800 CTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 1859

QY 778 ----- 777
Db 1860 TCACACAGATGTTTATCAACAACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGA 1919
QY 778 ----- 777
Db 1920 AGGTTCCGATGATGCAGTCTCTGTACAAAAGACGTTTGGATAACATGAACCTCAAGTGGAG 1979
QY 778 -----CAGTTCTGACCAGTG 792
Db 1980 TGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTTGGAAAGCCAGTTCTTGACCAGTG 2039
QY 793 GAAGCGTCTGCACCTTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAAGATGATGA 852
Db |||||
QY 2040 GAAGCGTCTGCACCTTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAAGATGATGA 2099
QY 853 ATTAAGCCGGCAGGCACCTATTGGAGCGGACCTTTCCAGCAGTTCCAGAAAGCAGAACGATGT 912
Db |||||
QY 2100 ATTAAGCCGGCAGGCACCTATTGGAGCGGACCTTTCCAGCAGTTTCCAGAAAGCAGAACGATGT 2159
QY 913 ACATAGGGCCCTCAAGAGGGGAATTGAAAACTAAAGAACCTGTATATCATGACTCTTTGA 972
Db |||||
QY 2160 ACATAGGGCCCTCAAGAGGGGAATTGAAAACTAAAGAACCTGTATATCATGACTCTTTGA 2219
QY 973 GACTGTACGAATATTTCTGACAGAGCGCCTTTTGGAAAGGACTAGAGAAACTCTACCAGGA 1032
Db |||||
QY 2220 GACTGTACGAATATTTCTGACAGAGCGCCTTTTGGAAAGGACTAGAGAAACTCTACCAGGA 2279
QY 1033 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGGCCCAAGAAATGTCACTCGGCTTCTACGAAAGCA 1092
Db |||||
QY 2280 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGGCCCAAGAAATGTCACTCGGCTTCTACGAAAGCA 2339
QY 1093 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGAACCTGCACTCCGCTGACTGGCAGAG 1152
Db |||||
QY 2340 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGAACCTGCACTCCGCTGACTGGCAGAG 2399
QY 1153 AAAAATAGATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 1212
Db |||||
QY 2400 AAAAATAGATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 2459
QY 1213 CCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCGCTGGGCGATCTCCT 1272
Db |||||
QY 2460 CCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCGCTGGGCGATCTCCT 2519
QY 1273 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAAAATTCGCC 1332
Db |||||
QY 2520 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAAAATTCGCC 2579
QY 1333 TCTGAAAGAGAACGTCAGGCCACGTCATGACCTTGTCTGCCAGCTTTACCACCTTTGGGCAT 1392
Db |||||
QY 2580 TCTGAAAGAGAACGTCAGGCCACGTCATGACCTTGTCTGCCAGCTTTACCACCTTTGGGCAT 2639
QY 1393 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATGGAAGCTTCT 1452
Db |||||
QY 2640 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATGGAAGCTTCT 2699
QY 1453 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGACCTTTGGTCC 1512
Db |||||
QY 2700 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGACCTTTGGTCC 2759
QY 1513 AGCATCTCAGCACTTTCTTTTCCACGTCCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCC 1572
Db |||||
QY 2760 AGCATCTCAGCACTTTCTTTTCCACGTCCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCC 2819
QY 1573 AAACAAAGTGCCCTACTATATCAACCAGAGACTCAACAACTTGTCTGGGACCATCCCAA 1632
Db |||||
QY 2820 AAACAAAGTGCCCTACTATATCAACCAGAGACTCAACAACTTGTCTGGGACCATCCCAA 2879
QY 1633 AATGACAGAGCTCTACAGTCTTTTAGCTGACCTGAATAATGTGAGATTCTCAGCTTATAG 1692
Db |||||
QY 2880 AATGACAGAGCTCTACAGTCTTTTAGCTGACCTGAATAATGTGAGATTCTCAGCTTATAG 2939

QY 778 ----- 777
Db 2557 CTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 2616
QY 778 ----- 777
Db 2617 TCACACAGATGTTTATCAACAACCTGGATGAAAAACAGCCAAAAATCCTGAGATCCCTGGA 2676
QY 778 ----- 777
Db 2677 AGGTTCCGATGATGCAGTCTCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAG 2736
QY 778 ----- CAGTTCTGACCAAGTG 792
Db 2737 TGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAAGCCAGTTCTGACCAAGTG 2796
QY 793 GAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGA 852
Db 2797 GAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGA 2856
QY 853 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTTCCAGAAAGCAGAAACGATGT 912
Db 2857 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTTCCAGAAAGCAGAAACGATGT 2916
QY 913 ACATAGGCGCTTCAAGAGGGGAATTGAAACTAAAGAACCTGTATCATGAGTACTCTTGA 972
Db 2917 ACATAGGCGCTTCAAGAGGGGAATTGAAACTAAAGAACCTGTATCATGAGTACTCTTGA 2976
QY 973 GACTGTACGAATATTTCTGACAGAGCAGCCTTTTGGAGGACTAGAGAAACTCTACCAGGA 1032
Db 2977 GACTGTACGAATATTTCTGACAGAGCAGCCTTTTGGAGGACTAGAGAAACTCTACCAGGA 3036
QY 1033 GCCCAGAGAGCTGCTCCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCA 1092
Db 3037 GCCCAGAGAGCTGCTCCTGAGGAGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCA 3096
QY 1093 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCGGCTGACTGGCAGAG 1152
Db 3097 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCGGCTGACTGGCAGAG 3156
QY 1153 AAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGA 1212
Db 3157 AAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGA 3216
QY 1213 CCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCTGGCGGATCTCCT 1272
Db 3217 CCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCTGGCGGATCTCCT 3276
QY 1273 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCCAGGAGAAATTGGGCC 1332
Db 3277 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCCAGGAGAAATTGGGCC 3336
QY 1333 TCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAT 1392
Db 3337 TCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAT 3396
QY 1393 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCT 1452
Db 3397 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCT 3456
QY 1453 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGGACTTTGGTCC 1512
Db 3457 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGGACTTTGGTCC 3516
QY 1513 AGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 1572
Db 3517 AGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 3576
QY 1573 AAACAAAGTCCCTACTATATCAACCACGAGACTCAACAACTTGCTGGGACCATCCCAA 1632
Db 3577 AAACAAAGTCCCTACTATATCAACCACGAGACTCAACAACTTGCTGGGACCATCCCAA 3636
QY 1633 AATGACAGAGCTCTACCAAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 1692

Db 3637 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 3696
QY 1693 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 1752
Db 3697 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 3756
QY 1753 AGCTGCATGTGATGCTTGGACCAAGCAAACTCAAGCAAAATGACAGCCCATGGATAT 1812
Db 3757 AGCTGCATGTGATGCTTGGACCAAGCAAACTCAAGCAAAATGACAGCCCATGGATAT 3816
QY 1813 CCTGCAGATTATTAATTGTTTGACCACTATTTATGACCGCCTGGAGCAAGAGCACAACAA 1872
Db 3817 CCTGCAGATTATTAATTGTTTGACCACTATTTATGACCGCCTGGAGCAAGAGCACAACAA 3876
QY 1873 TTTGGTCAACGTCCTCTCTCGTGGATATGTCTGAACCTGGCTGCTGAATGTTTATGA 1932
Db 3877 TTTGGTCAACGTCCTCTCTCGTGGATATGTCTGAACCTGGCTGCTGAATGTTTATGA 3936
QY 1933 TACGGGACGAACAGGAGGATCCGTGTCCTGCTTTTAAAACTGGCATCATTTCCCTGTG 1992
Db 3937 TACGGGACGAACAGGAGGATCCGTGTCCTGCTTTTAAAACTGGCATCATTTCCCTGTG 3996
QY 1993 TAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 2052
Db 3997 TAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 4056
QY 2053 ATTTTGTGACCAAGCGCAGGCTGGGCTCCTTCTGTCATGATTCTATCCAA 2101
Db 4057 ATTTTGTGACCAAGCGCAGGCTGGGCTCCTTCTGTCATGATTCTATCCAA 4105

RESULT 8
AAD37263

ID AAD37263 standard; DNA; 4848 BP.

XX AAD37263;

XX 21-AUG-2002 (first entry)

XX Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3849.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.

XX Chimeric - Homo sapiens.

OS Chimeric - Cytomegalovirus.

OS Chimeric - Unidentified.

XX WO200183695-A2.

XX 08-NOV-2001.

XX 27-APR-2001; 2001WO-US13677.

XX 28-APR-2000; 2000US-200777P.

XX (XIAO/) XIAO X.

XX xiao X;

XX WPI; 2002-049342/06.

XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -

XX Example 1; Page 68-70; 71pp; English.

XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified

CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC cytomagalovirus (CMV) promoter and a small polyA signal sequence.
XX
SQ Sequence 4848 BP; 1405 A; 1144 C; 1180 G; 1119 T; 0 other;

Query Match 83.0%; Score 1743; DB 24; Length 4848;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 348; Gaps 1;

| | | | |
|----|------|--|------|
| QY | 1 | GAGCTATGCTACACACAGGCTGCTTATGTCAACCCTCTGACCCCTACACGGAGCCCAT | 60 |
| Db | 1680 | GAGCTATGCTACACACAGGCTGCTTATGTCAACCCTCTGACCCCTACACGGAGCCCAT | 1739 |
| QY | 61 | TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTTCATTGATGAGAG | 120 |
| Db | 1740 | TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTTCATTGATGAGAG | 1799 |
| QY | 121 | TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAGATTTATCGTGGCTTCTTC | 180 |
| Db | 1800 | TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAGATTTATCGTGGCTTCTTC | 1859 |
| QY | 181 | TGCTGAGGACACATTCAGACCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA | 240 |
| Db | 1860 | TGCTGAGGACACATTCAGACCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA | 1919 |
| QY | 241 | CCAGTTTCATCTACATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG | 300 |
| Db | 1920 | CCAGTTTCATCTACATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG | 1979 |
| QY | 301 | TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA | 360 |
| Db | 1980 | TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA | 2039 |
| QY | 361 | AACTGAAGTACAAGACGATGAATCTCTTAATTTCAAGATGGGAATGCCTCAGGGTAGC | 420 |
| Db | 2040 | AACTGAAGTACAAGACGATGAATCTCTTAATTTCAAGATGGGAATGCCTCAGGGTAGC | 2099 |
| QY | 421 | TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT | 480 |
| Db | 2100 | TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT | 2159 |
| QY | 481 | GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGGAAATGGAGGAAGA | 540 |
| Db | 2160 | GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGGAAATGGAGGAAGA | 2219 |
| QY | 541 | GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAACAATAGGTGCTTCA | 600 |
| Db | 2220 | GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAACAATAGGTGCTTCA | 2279 |
| QY | 601 | AGAAGATCTAGAACACAGAAACAGTCAAGGTCAAFTCTCTCACTCACATGGTGGTAGT | 660 |
| Db | 2280 | AGAAGATCTAGAACACAGAAACAGTCAAGGTCAAFTCTCTCACTCACATGGTGGTAGT | 2339 |
| QY | 661 | TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG | 720 |
| Db | 2340 | TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG | 2399 |
| QY | 721 | AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGA --- | 777 |
| Db | 2400 | AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAC | 2459 |
| QY | 778 | ----- | 777 |
| Db | 2460 | TCATAGATTACTGCAACAGTTCCTCCCTGGACCTTGGAAGAGTTTCTTGCTGGGCTTACAGA | 2519 |

| | | | |
|----|------|---|------|
| QY | 778 | ----- | 777 |
| Db | 2520 | AGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAAGGCTCTTAGAAGA | 2579 |
| QY | 778 | ----- | 777 |
| Db | 2580 | CTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAATTAAGC | 2639 |
| QY | 778 | ----- | 777 |
| Db | 2640 | TCACACAGATGTTTATCACAACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGA | 2699 |
| QY | 778 | ----- | 777 |
| Db | 2700 | AGGTTCCGATGATGAGTCCTGTTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAG | 2759 |
| QY | 778 | ----- | 792 |
| Db | 2760 | TGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTG | 2819 |
| QY | 793 | GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA | 852 |
| Db | 2820 | GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA | 2879 |
| QY | 853 | ATTAAGCCGGCAGGCACCTATTGGAGGCGGACTTTCCAGCAGTTTCAAGAGCAGAACCATGT | 912 |
| Db | 2880 | ATTAAGCCGGCAGGCACCTATTGGAGGCGGACTTTCCAGCAGTTTCAAGAGCAGAACCATGT | 2939 |
| QY | 913 | ACATAGGGCCTTCAAGAGGGAATTTAAAAAATAAAGAACCTGTAATCATGAGTACTTTGA | 972 |
| Db | 2940 | ACATAGGGCCTTCAAGAGGGAATTTAAAAAATAAAGAACCTGTAATCATGAGTACTTTGA | 2999 |
| QY | 973 | GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGA | 1032 |
| Db | 3000 | GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGA | 3059 |
| QY | 1033 | CCCCAGAGAGCTCCTCCTGAGGAGAGAGCCAGAAATGTCACCTCGGCTTCTACGAAGCA | 1092 |
| Db | 3060 | CCCCAGAGAGCTCCTCCTGAGGAGAGAGCCAGAAATGTCACCTCGGCTTCTACGAAGCA | 3119 |
| QY | 1093 | GGCTGAGGAGGTCAATACTAGTGGGAAAAATTTGAACCTGCACCTCCGCTGACTGGCAGAG | 1152 |
| Db | 3120 | GGCTGAGGAGGTCAATACTAGTGGGAAAAATTTGAACCTGCACCTCCGCTGACTGGCAGAG | 3179 |
| QY | 1153 | AAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA | 1212 |
| Db | 3180 | AAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA | 3239 |
| QY | 1213 | CCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCT | 1272 |
| Db | 3240 | CCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCT | 3299 |
| QY | 1273 | CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACITTCGAGGAGAAATTTGCGCC | 1332 |
| Db | 3300 | CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACITTCGAGGAGAAATTTGCGCC | 3359 |
| QY | 1333 | TCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAT | 1392 |
| Db | 3360 | TCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAT | 3419 |
| QY | 1393 | TCAGCTCTCACCCTGATAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCT | 1452 |
| Db | 3420 | TCAGCTCTCACCCTGATAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCT | 3479 |
| QY | 1453 | GCAGGTGGCCGTCGAGGACCGAGTCAGGTCAGGTGATGAAGCCACAGGGACTTTTGGTCC | 1512 |
| Db | 3480 | GCAGGTGGCCGTCGAGGACCGAGTCAGGTCAGGTGATGAAGCCACAGGGACTTTTGGTCC | 3539 |
| QY | 1513 | AGCATCTCAGCACCTTTCTTTCCAGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC | 1572 |
| Db | 3540 | AGCATCTCAGCACCTTTCTTTCCAGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC | 3599 |
| QY | 1573 | AAACAAAGTGCCCTACTATATATCAACCAACGAGACTCAAAACAACTTGTGGGACCATCCCAA | 1632 |

Db 3600 AAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACAACCTTGCTGGACCATCCCAA 3659
Qy 1633 AATGACAGAGCTACCAAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 1692
Db 3660 AATGACAGAGCTACCAAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 3719
Qy 1693 GACTGCCATGAACCTCCGAAGACTGCAGAGGCCCTTTGCTTGATCTCTTGAGCCCTGTC 1752
Db 3720 GACTGCCATGAACCTCCGAAGACTGCAGAGGCCCTTTGCTTGATCTCTTGAGCCCTGTC 3779
Qy 1753 AGCTGCATGTATGCTTGACCAAGCAGCACAACTCAAGCAAAATGACCAAGCCCATGGATAT 1812
Db 3780 AGCTGCATGTATGCTTGACCAAGCAGCACAACTCAAGCAAAATGACCAAGCCCATGGATAT 3839
Qy 1813 CCTGCAGATTATTAATGTTTGGACCACTATTATGACCGCCTGGAGCAAGAGCACAAACA 1872
Db 3840 CCTGCAGATTATTAATGTTTGGACCACTATTATGACCGCCTGGAGCAAGAGCACAAACA 3899
Qy 1873 TTTGGTCAACGTCCTCTCTGCGTGGATATGTTGTAACCTGGCTGCTGAATGTTTATGA 1932
Db 3900 TTTGGTCAACGTCCTCTCTGCGTGGATATGTTGTAACCTGGCTGCTGAATGTTTATGA 3959
Qy 1933 TACGGGACGAACAGGAGGATCCGTGTCCTGCTTTTAAACCTGGCATCATTTCCCTGTG 1992
Db 3960 TACGGGACGAACAGGAGGATCCGTGTCCTGCTTTTAAACCTGGCATCATTTCCCTGTG 4019
Qy 1993 TAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 2052
Db 4020 TAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 4079
Qy 2053 ATTTTGTGACCGAGGCTGGGCTCCCTTCTGTCATGATTCTATCCAA 2101
Db 4080 ATTTTGTGACCGAGGCTGGGCTCCCTTCTGTCATGATTCTATCCAA 4128

RESULT 9
AAD37264
ID AAD37264 standard; DNA; 5060 BP.
XX AAD37264;
AC
XX 21-AUG-2002 (first entry)
DT
XX
DE Adeno-associated virus (AAV) vector plasmid, AAV-E-CMV-3849.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Cytomegalovirus.
OS Chimeric - Unidentified.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX

PS Example 1; Page 70-71; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC muscle creatine kinase (MCK) enhancer, a cytomegalovirus (CMV) promoter
CC and a small polyA signal sequence.
XX
SQ Sequence 5060 BP; 1449 A; 1217 C; 1234 G; 1160 T; 0 other;
Query Match 83.0%; Score 1743; DB 24; Length 5060;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 348; Gaps 1;
Qy 1 GAGCTATGCCTACACACAGGCTGCTTATGTACCAACCTCTGACCCCTACACGGAGCCCAT 60
Db 1892 GAGCTATGCCTACACACAGGCTGCTTATGTACCAACCTCTGACCCCTACACGGAGCCCAT 1951
Qy 61 TCCTTCACAGCATTGGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 120
Db 1952 TCCTTCACAGCATTGGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 2011
Qy 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATTTATGCTGGCTTCTTTC 180
Db 2012 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATTTATGCTGGCTTCTTTC 2071
Qy 181 TGCTGAGGACACATTGCAAGCACCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 2072 TGCTGAGGACACATTGCAAGCACCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 2131
Qy 241 CCAGTTTTCATCTACTGAGGGGTACATGATGGATTGACAGCCCATCAGGCCGGTGTGG 300
Db 2132 CCAGTTTTCATCTACTGAGGGGTACATGATGGATTGACAGCCCATCAGGCCGGTGTGG 2191
Qy 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 360
Db 2192 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 2251
Qy 361 AACTGAAGTACAAGACAGATGAATCTCTTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
Db 2252 AACTGAAGTACAAGACAGATGAATCTCTTAAATTCAGATGGGAATGCCTCAGGGTAGC 2311
Qy 421 TAGCATGGAAAAACAAAGCAATTTACATAGATTTTAATGGATCTCCAGAAATCAGAAACT 480
Db 2312 TAGCATGGAAAAACAAAGCAATTTACATAGATTTTAATGGATCTCCAGAAATCAGAAACT 2371
Qy 481 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAACAAGGAAATGGAGGAAGA 540
Db 2372 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAACAAGGAAATGGAGGAAGA 2431
Qy 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAATAGGTGCTTCA 600
Db 2432 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAATAGGTGCTTCA 2491
Qy 601 AGAAGATCTAGAACACAGAACCAAGTCAGGGTCAATTTCTCACTCACATGGTGGTAGT 660
Db 2492 AGAAGATCTAGAACACAGAACCAAGTCAGGGTCAATTTCTCACTCACATGGTGGTAGT 2551
Qy 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGTTTGAAGAAACAACATTAAGGTATTGGG 720
Db 2552 TGATGAATCTAGTGGAGATCACGCAACTGCTGTTTGAAGAAACAACATTAAGGTATTGGG 2611
Qy 721 AGATCGATGGCAACACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGA --- 777
Db 2612 AGATCGATGGCAACACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAC 2671

QY 778 ----- 777
Db 2672 TCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAGTTTCTTGCTGGCTTACAGA 2731
QY 778 ----- 777
Db 2732 AGCTGAAACAACACTGCCAATGTCTACAGGATGCTACCCGTAAAGAAAGCTCCTAGAAGA 2791
QY 778 ----- 777
Db 2792 CTCCAAGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGAATTGAAGC 2851
QY 778 ----- 777
Db 2852 TCACACAGATGTTTATCACAAACCTGGATCAAAAACAGCCAAAATAATCCTGAGATCCCTGGA 2911
QY 778 ----- 777
Db 2912 AGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAG 2971
QY 778 ----- CAGTTCTGACCAGTG 792
Db 2972 TGAAC TTCGGA AAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTG 3031
QY 793 GAAGCGTCTGCACCTTTCTCTGCAGGAACCTCTGGTGTGGCTACAGCTGAAAGATGATGA 852
Db 3032 GAAGCGTCTGCACCTTTCTCTGCAGGAACCTCTGGTGTGGCTACAGCTGAAAGATGATGA 3091
QY 853 ATTAAGCCGGCAGGCACCTATTGGAGCGGACCTTTCAGCAGTTTCAGAAGCAGAACGATGT 912
Db 3092 ATTAAGCCGGCAGGCACCTATTGGAGCGGACCTTTCAGCAGTTTCAGAAGCAGAACGATGT 3151
QY 913 ACATAGGGCCTTCAAGAGGGAATTGAAAACCTPAAAGAACCTGTATCATGACTCTTTGA 972
Db 3152 ACATAGGGCCTTCAAGAGGGAATTGAAAACCTPAAAGAACCTGTATCATGACTCTTTGA 3211
QY 973 GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACTCTACCAGGA 1032
Db 3212 GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACTCTACCAGGA 3271
QY 1033 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCA 1092
Db 3272 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCA 3331
QY 1093 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAG 1152
Db 3332 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAG 3391
QY 1153 AAAATPAGATGAGACCCCTTGAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGA 1212
Db 3392 AAAATPAGATGAGACCCCTTGAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGA 3451
QY 1213 CCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCT 1272
Db 3452 CCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCT 3511
QY 1273 CATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACTTCGAGGAGAAATTGCGCC 1332
Db 3512 CATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACTTCGAGGAGAAATTGCGCC 3571
QY 1333 TCTGAAAGAGAACGCTGAGCCACGTCAAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAT 1392
Db 3572 TCTGAAAGAGAACGCTGAGCCACGTCAAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAT 3631
QY 1393 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTTGAACACCAGATGGAGCTTCT 1452
Db 3632 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTTGAACACCAGATGGAGCTTCT 3691
QY 1453 GCAGGTGGCCGCTCGAGGACCCGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTGGTCC 1512
Db 3692 GCAGGTGGCCGCTCGAGGACCCGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTGGTCC 3751

QY 1513 AGCATCTCAGCACTTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 1572
Db 3752 AGCATCTCAGCACTTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 3811
QY 1573 AAACAAAGTGCCTACTATATCAACCAACGAGACTCAACAACTTGTCTGGGACCATCCCAA 1632
Db 3812 AAACAAAGTGCCTACTATATCAACCAACGAGACTCAACAACTTGTCTGGGACCATCCCAA 3871
QY 1633 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCTCAGCTTATAG 1692
Db 3872 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCTCAGCTTATAG 3931
QY 1693 GACTGCCATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTG 1752
Db 3932 GACTGCCATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTG 3991
QY 1753 AGCTGCATGTGATGCCCTTGGACCAACCAACCTCAAGCAAAATGACCAAGCCCATGGATAT 1812
Db 3992 AGCTGCATGTGATGCCCTTGGACCAACCAACCTCAAGCAAAATGACCAAGCCCATGGATAT 4051
QY 1813 CCTGCAGATTATTAATTTGACCACTATTTATGACCGCCTGGAGCAAGACACAA 1872
Db 4052 CCTGCAGATTATTAATTTGACCACTATTTATGACCGCCTGGAGCAAGACACAA 4111
QY 1873 TTTGGTCAACGTCCTCTCTGCGTGGATATGTCTGAACCTGGCTGTGAATGTTTATGA 1932
Db 4112 TTTGGTCAACGTCCTCTCTGCGTGGATATGTCTGAACCTGGCTGTGAATGTTTATGA 4171
QY 1933 TACGGGACGAACAGGAGGATCCGTGTCTCTTTTAAACCTGGCATCATTTCCCTGTG 1992
Db 4172 TACGGGACGAACAGGAGGATCCGTGTCTCTTTTAAACCTGGCATCATTTCCCTGTG 4231
QY 1993 TAAAGCACATTTTGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 2052
Db 4232 TAAAGCACATTTTGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 4291
QY 2053 ATTTTGTGACCGCAGGCTGGGCTCCTTCTGTCATGATTCTATCCAA 2101
Db 4292 ATTTTGTGACCGCAGGCTGGGCTCCTTCTGTCATGATTCTATCCAA 4340

RESULT 10

AAD37260
ID AAD37260 standard; DNA; 4414 BP.

XX AAD37260;

XX 21-AUG-2002 (first entry)

XX Adeno-associated virus vector plasmid, AAV-MCK-3447.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.

XX Chimeric - Homo sapiens.
OS Chimeric - Unidentified.

XX WO200183695-A2.

XX 08-NOV-2001.

XX 27-APR-2001; 2001WO-US13677.

XX 28-APR-2000; 2000US-200777P.

XX (XIAO/) XIAO X.

XX Xiao X;

XX WPI; 2002-049342/06.

XX New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
XX dystrophin gene -
XX
PS Example 1; Page 65-66; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence.
XX
SQ Sequence 4414 BP; 1255 A; 1075 C; 1086 G; 998 T; 0 other;

Query Match 78.1%; Score 1641; DB 24; Length 4414;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1848; Conservative 0; Mismatches 190; Indels 63; Gaps 3;

| | | | |
|----|------|--|-----|
| QY | 1 | GAGCTATGCCTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCAT | 60 |
| DB | 1657 | | |
| QY | 61 | TCCTTACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCAATTGATGGAGAG | 120 |
| DB | 1717 | | |
| QY | 121 | TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC | 180 |
| DB | 1777 | | |
| QY | 181 | TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA | 240 |
| DB | 1837 | | |
| QY | 241 | CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG | 300 |
| DB | 1897 | | |
| QY | 301 | TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTTATCAGAAGATGAAGA | 360 |
| DB | 1957 | | |
| QY | 361 | AAGTGAAGTACAAGACAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC | 420 |
| DB | 2017 | | |
| QY | 421 | TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT | 480 |
| DB | 2077 | | |
| QY | 481 | GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAGAAAGAACCAAGGAAAATGGAGGAAGA | 540 |
| DB | 2137 | | |
| QY | 541 | GCCTCTTGACCTGATCTTGAAGACCTAAACGCGCAAGTACAAACAATAAGGTGCTTCA | 600 |
| DB | 2197 | | |
| QY | 601 | AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTTCTCACTCACTACATGGTGGTAGT | 660 |
| DB | 2257 | | |
| QY | 661 | TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG | 720 |
| DB | 2317 | | |

| | | | |
|----|------|---|------|
| QY | 721 | AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGTCTTTTACAAGACAG | 780 |
| DB | 2377 | | |
| QY | 781 | TTCTGACCAAGTGGGAAGCGTCTGCACCTTTCTCTGACGGAACCTTCTGGTGTGGCTACAGCT | 840 |
| DB | 2437 | | |
| QY | 841 | GAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAA | 900 |
| DB | 2497 | | |
| QY | 901 | GCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTTGAAAACTAAAGAACCTGTAATCAT | 960 |
| DB | 2557 | | |
| QY | 961 | GAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAA | 1020 |
| DB | 2617 | | |
| QY | 1021 | ACTCTACAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCT | 1080 |
| DB | 2663 | | |
| QY | 1081 | TCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGAACCTGCACCTCCGC | 1140 |
| DB | 2721 | | |
| QY | 1141 | TGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCAC | 1200 |
| DB | 2744 | | |
| QY | 1201 | GGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT | 1260 |
| DB | 2794 | | |
| QY | 1261 | GGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGG | 1320 |
| DB | 2854 | | |
| QY | 1321 | AGAAATTGGCCTCTGAAAGAGAACGTTGAGCCACGTCATGACCTTGCTGCCAGCTTAC | 1380 |
| DB | 2914 | | |
| QY | 1381 | CACTTTGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAG | 1440 |
| DB | 2974 | | |
| QY | 1441 | ATGGAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAG | 1500 |
| DB | 3034 | | |
| QY | 1501 | GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTGTCCAGGGTCCCTGGGAGAG | 1560 |
| DB | 3094 | | |
| QY | 1561 | AGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACTTGGCTG | 1620 |
| DB | 3154 | | |
| QY | 1621 | GGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAATGTGAGATT | 1680 |
| DB | 3214 | | |
| QY | 1681 | CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCT | 1740 |
| DB | 3274 | | |
| QY | 1741 | CTTGAGCCTGTGAGCTGCATGTGATGCCCTTGAGCCAGCACAACCTCAAGCAAAATGACCA | 1800 |
| DB | 3334 | | |
| QY | 1801 | GCCCATGGATATCCTGCGAGATTATTAAATGTTTGACCACCTATTTATGACCGCCTGGAGCA | 1860 |

```
Db 3394 GCCCATGGATATCCTGCGAGATTATTAATGTTTGGACCACTATTATGACCGCCTGGAGCA 3453
QY 1861 AGAGCACAAACAAATTTGGTCAACGTCCTCTCTGCGTGGATATGTGTGAACCTGGCTGCT 1920
Db 3454 AGAGCACAAACAAATTTGGTCAACGTCCTCTCTGCGTGGATATGTGTGAACCTGGCTGCT 3513
QY 1921 GAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGTCCTTTTAAAACTGGCAT 1980
Db 3514 GAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGTCCTTTTAAAACTGGCAT 3573
QY 1981 CATTTCCCTGTGTAAGCACATTTGGAGACAAGTACAGATACCTTTTCAAGCAAAGTGGC 2040
Db 3574 CATTTCCCTGTGTAAGCACATTTGGAGACAAGTACAGATACCTTTTCAAGCAAAGTGGC 3633
QY 2041 AAGTTCAACAGGATTTGTGACCAGCGCGAGGCTGGGCTCCTTCGTCATGATTCTATCCA 2100
Db 3634 AAGTTCAACAGGATTTGTGACCAGCGCGAGGCTGGGCTCCTTCGTCATGATTCTATCCA 3693
QY 2101 A 2101
Db 3694 A 3694

RESULT 11
AAD37242
ID AAD37242 standard; DNA; 3446 BP.
XX
AC AAD37242;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta3447.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 53-54; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3447 containing nucleotides 1-1992 (N-terminus,
CC hinge H1 and rods R1, R2 and R3), 8749-10227 (rod R24, hinge H4 and
```

```
CC CR domain) and 11047-11058 (dystrophin last 3 amino acids).
XX
SQ Sequence 3446 BP; 1074 A; 766 C; 787 G; 819 T; 0 other;
Query Match 77.5%; Score 1629; DB 24; Length 3446;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 1847; Conservative 0; Mismatches 190; Indels 64; Gaps 4;
QY 1 GAGCTATGCCTACACACAGGGTGTGTTATGTCAACCCTCTGACCCCTACACGGAGCCCAT 60
Db 900 GAGCTATGCCTACACACAGGGTGTGTTATGTCAACCCTCTGACCCCTACACGGAGCCCAT 959
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 120
Db 960 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 1019
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
Db 1020 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1079
QY 181 TGCTGAGGACACATTTCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1080 TGCTGAGGACACATTTCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1139
QY 241 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG 300
Db 1140 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG 1199
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTTATCAGAAGATGAAGA 360
Db 1200 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTTATCAGAAGATGAAGA 1259
QY 361 AACTGAAGTACAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
Db 1260 AACTGAAGTACAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 1319
QY 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 480
Db 1320 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATC-GAAACT 1378
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAACAAGGAAAAATGGAGGAAGA 540
Db 1379 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAACAAGGAAAAATGGAGGAAGA 1438
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAATAAGGTGCTTCA 600
Db 1439 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACAATAAGGTGCTTCA 1498
QY 601 AGAAGATCTAGAACAAACAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGTAGT 660
Db 1499 AGAAGATCTAGAACAAACAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGTAGT 1558
QY 661 TGATGAATCTAGTGGAGATCAGCAACTGTGCTTTTGGGAAGACAACCTTAAGGTATTGGG 720
Db 1559 TGATGAATCTAGTGGAGATCAGCAACTGTGCTTTTGGGAAGACAACCTTAAGGTATTGGG 1618
QY 721 AGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGGGTTCTTTTACAAGACAG 780
Db 1619 AGATCGATGGGCAACATCTGTAGATGGACAGAACCCGCTGGGTTCTTTTACAAGACAT 1678
QY 781 TTCTGACCAAGTGGAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCT 840
Db 1679 CCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTATGTCATGGCTTTCAGA 1738
QY 841 GAAAGATGATGAATTAAGCCGGCAGGCACCTTAFTGGAGCGCACTTTCCAGCAGTTTCAGAA 900
Db 1739 AAAAGAAGATGCAGTGAACAAGATTCACACAACCTGGCTTTTAAAGATCAAAATGAATGTT 1798
QY 901 GCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTTGAAAACTAAAGAACTGTATCAT 960
Db 1799 ATCAGTCTTCAAAAACTGGCCGTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 1858
QY 961 GAGTACTCTTGAGACTGTACGAATATTTCTGCAGAGCAGCCCTTTGGGAAGGACTAGAGAA 1020
```


Db 1859 GGGCAAACTGTA-----TTCACTCAAAACAAGATCTTCTTTCAACACTGAAG 1904

Qy 1021 ACTCTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAAGATGTCACTCGGCT 1080

Db 1905 AATAAGTCAGTGACCCAGAAGACGGAAGCATGCGTGGA----- 1942

Qy 1081 TCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCACTCCGC 1140

Db 1943 -----TAACTTTGCCCGGTGTGGGATAATTTAGTCCAAAAAACTTGAA----- 1985

Qy 1141 TGA CTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCAC 1200

Db 1986 -----AAGAGTACAGCACAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCAC 2035

Qy 1201 GGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT 1260

Db 2036 GGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT 2095

Qy 1261 GGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGG 1320

Db 2096 GGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGG 2155

Qy 1321 AGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTAC 1380

Db 2156 AGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTAC 2215

Qy 1381 CACTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGGAACACCCAG 1440

Db 2216 CACTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGGAACACCCAG 2275

Qy 1441 ATGGAAGCTTCTGCAGGTGGCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAG 1500

Db 2276 ATGGAAGCTTCTGCAGGTGGCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAG 2335

Qy 1501 GGACTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTCGTCCAGGGTCCCTGGGAGAG 1560

Db 2336 GGACTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTCGTCCAGGGTCCCTGGGAGAG 2395

Qy 1561 AGCCATCTGCCAAACAAAGTGCCTACTATATCAACCAGAGACTCAAAACAACCTTGCTG 1620

Db 2396 AGCCATCTGCCAAACAAAGTGCCTACTATATCAACCAGAGACTCAAAACAACCTTGCTG 2455

Qy 1621 GGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTTAGTGACCTGAATAATGTCAGATT 1680

Db 2456 GGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTTAGTGACCTGAATAATGTCAGATT 2515

Qy 1681 CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTGGATCT 1740

Db 2516 CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTGGATCT 2575

Qy 1741 CTTGAGCCTGTGAGCTGCATGTGATGCTTGGACCGACCAACCTCAAGCAAAATGACCA 1800

Db 2576 CTTGAGCCTGTGAGCTGCATGTGATGCTTGGACCGACCAACCTCAAGCAAAATGACCA 2635

Qy 1801 GCCCATGGATATCCTGCAGATTATTAATTGTTTGGACCACTATTTATGACCGCCTGGAGCA 1860

Db 2636 GCCCATGGATATCCTGCAGATTATTAATTGTTTGGACCACTATTTATGACCGCCTGGAGCA 2695

Qy 1861 AGAGCACAAACAATTTGGTCAACGTCCTCTCTGCGTGGATATGTGTCTGAACCTGGCTGCT 1920

Db 2696 AGAGCACAAACAATTTGGTCAACGTCCTCTCTGCGTGGATATGTGTCTGAACCTGGCTGCT 2755

Qy 1921 GAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGCTTTTAAAACTGGCAT 1980

Db 2756 GAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGCTTTTAAAACTGGCAT 2815

Qy 1981 CATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGC 2040

Db 2816 CATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGC 2875

Qy 2041 AAGTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGGCTCCTCTCTGTCATGATTCTATCCA 2100

Db 2876 AAGTTCAACAGGATTTGTGACCAGCGCAGGCTGGGCTCCTTCTGCAATGATTCTATCCA 2935

Qy 2101 A 2101

Db 2936 A 2936

RESULT 12

ABK81999

ID ABK81999 standard; DNA; 5462 BP.

XX

AC ABK81999;

XX

DT 13-AUG-2002 (first entry)

XX

DE DNA encoding mini-dystrophin protein deltaR2-R21+H3.

XX

KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;

KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.

XX

OS Homo sapiens.

OS Synthetic.

XX

PN WO200229056-A2.

XX

PD 11-APR-2002.

XX

PF 04-OCT-2001; 2001WO-US311126.

XX

PR 06-OCT-2000; 2000US-238848P.

XX

PA (UNMI) UNIV MICHIGAN.

XX

PI Chamberlain JS, Harper SQ;

XX

DR WPI; 2002-435334/46.

XX

PT A composition for preparing therapeutic drugs, has a mini-dystrophin peptide comprising a specific number of spectrin-like repeat domains, or a nucleic acid sequence encoding the mini-dystrophin peptide -

PS Disclosure; Fig 14; 145pp; English.

XX

CC The invention describes a composition comprising a mini-dystrophin peptide comprising a spectrin-like repeat domain, where the domain comprises n spectrin-like repeats, and contains no more than n spectrin-like repeats, where n is an even number between 4-24, or a nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the polynucleotide encoding it is useful as a medicament, for preparing a drug for therapeutic application and in the preparation of a composition for treatment of muscle disease, e.g. Duchenne's muscular dystrophy (DMD). This sequence represents a mini-dystrophin sequence of the invention.

XX

SQ Sequence 5462 BP; 1668 A; 1225 C; 1212 G; 1357 T; 0 other;

Query Match 77.5%; Score 1629; DB 24; Length 5462;

Best Local Similarity 85.4%; Pred. No. 0;

Matches 1911; Conservative 0; Mismatches 190; Indels 138; Gaps 3;

Qy 1 GAGCTATGCTTACACACAGGCTGCTTATGTCAACCCTGACCCCTACACGGAGCCCATTT 60

Db 1099 GAGCTATGCTTACACACAGGCTGCTTATGTCAACCCTGACCCCTACACGGAGCCCATTT 1158

Qy 61 TCCTTCACAGCATTTGGAAGCTCTGGAAGCAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120

Db 1159 TCCTTCACAGCATTTGGAAGCTCTGGAAGCAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1218

Qy 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 180

Db 1219 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 1278

Qy 181 TGCTGAGGACACATTTGCAAGCACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240

Db 1279 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGANTTTCTAATGATGTGGAAGTGGTGAAGA 1338
QY 241 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTGTACAGCCCATCAGGGCCGGGTTGG 300
Db 1339 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTGTACAGCCCATCAGGGCCGGGTTGG 1398
QY 301 TAAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1399 TAAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 1458
QY 361 AACTGAAGTACAAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518
QY 421 TAGCATGGAAAAACAAAGCAATTACATAGAGTT----- 454
Db 1519 TAGCATGGAAAAACAAAGCAATTACATGCTCCTGGACTGACCACTATTGGAGCCTCTCC 1578
QY 455 ----- 454
Db 1579 TACTCAGACTGTTACTCTCTGTGACACAACCTGTGTGTTACTAAGGAAACTGCCATCTCCAA 1638
QY 455 -----TTAATGGATCTCCAGAATCAGAA 477
Db 1639 ACTAGAAATGCCATCTTCTCTGATGTTGGAGCATAGATTACTGCAACAGTTCCTCCCTGGA 1698
QY 478 ACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAAACAAAGGAAAAATGGAGGA 537
Db 1699 CTTGGAAGAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGA 1758
QY 538 AGAGCCTCTTGGAACCTGATCTTGAAGACCTAAACGCCAAAGTACAAACAAACATAAGGTGCT 597
Db 1759 TGTACCCGTAAGGAAGGGCTCTAGAGACTCCAAGGGAGTAAAGAGCTGATGAACA 1818
QY 598 TCAAGAAATCTAGAACAAAGAAACAGTCAAGGTCAATTCTCTCACTCACATGTTGTTGGT 657
Db 1819 ATGGCAAGACCTCCAAGGTGAATTGAAGCTCACACAGATGTTTATCACAACTGGATGA 1878
QY 658 AGTTGATGA-----ATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACA 705
Db 1879 AAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTACAAAG 1938
QY 706 ACTTAAAG---TATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTG 762
Db 1939 ACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAG 1998
QY 763 GGTCTCTTTTACAAGACAGATTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAAGAACT 822
Db 1999 GTCCCATTTTGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAAGAACT 2058
QY 823 TCTGTTGGCTACAGCTGAAAGATGAATGAATTAAGCCGGCAGGCACCTATTGGAGGCGA 882
Db 2059 TCTGTTGGCTACAGCTGAAAGATGAATGAATTAAGCCGGCAGGCACCTATTGGAGGCGA 2118
QY 883 CTTTCCAGCAGTTCAGAAGCAGAACCGATGTACATAGGGCCTTCAAGAGGGAATTGAAAAAC 942
Db 2119 CTTTCCAGCAGTTCAGAAGCAGAACCGATGTACATAGGGCCTTCAAGAGGGAATTGAAAAAC 2178
QY 943 TAAAGAACCTCTAATCATGAGTACTCTTGAGACTGTACGAAATATTTCTGACAGAGCAGCC 1002
Db 2179 TAAAGAACCTCTAATCATGAGTACTCTTGAGACTGTACGAAATATTTCTGACAGAGCAGCC 2238
QY 1003 TTTGGAAGGACTAGAGAAAACTCTACCAGAGCCAGAGAGCTGCTCCTGAGGAGAGAGC 1062
Db 2239 TTTGGAAGGACTAGAGAAAACTCTACCAGAGCCAGAGAGCTGCTCCTGAGGAGAGAGC 2298
QY 1063 CCAGAAATGTCACCTCGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAA 1122
Db 2299 CCAGAAATGTCACCTCGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAA 2358
QY 1123 ATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAAATAGATGAGACCTTGAAGACTCCA 1182

Db 2359 ATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCA 2418
QY 1183 GGAACCTTCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGGCCAAAGCTGAGGTGATCAA 1242
Db 2419 GGAACCTTCAAGAGGCCACCGGATGAGCTGGACCTCAAGCTGGCCAAAGCTGAGGTGATCAA 2478
QY 1243 GGGATCCTGGCAGCCCGTGGCGGATCTCTCTATTGACTCTCTCCAAGATCACCTCGAGAA 1302
Db 2479 GGGATCCTGGCAGCCCGTGGCGGATCTCTCTATTGACTCTCTCCAAGATCACCTCGAGAA 2538
QY 1303 AGTCAAGGCACCTTCGAGGAGAAAAATTGGCGCTCTGAAAGAGAACGTGAGCCACGTCAATGA 1362
Db 2539 AGTCAAGGCACCTTCGAGGAGAAAAATTGGCGCTCTGAAAGAGAACGTGAGCCACGTCAATGA 2598
QY 1363 CTTTGCTCGCCAGTTTACCACCTTTGGGCAATTGAGCTCTCAACGATATAACCTCAGCACTCT 1422
Db 2599 CTTTGCTCGCCAGTTTACCACCTTTGGGCAATTGAGCTCTCAACGATATAACCTCAGCACTCT 2658
QY 1423 GGAAGACCTGAACACACAGATGGAAGCTTTGCGAGGTGGCCGTGAGGACCGAGTCAGGCA 1482
Db 2659 GGAAGACCTGAACACACAGATGGAAGCTTTGCGAGGTGGCCGTGAGGACCGAGTCAGGCA 2718
QY 1483 GCTGCATGAAGCCCAAGGAGACTTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTGT 1542
Db 2719 GCTGCATGAAGCCCAAGGAGACTTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTGT 2778
QY 1543 CCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAAAAAGTGCCCTACTATATCAACCCACGA 1602
Db 2779 CCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAAAAAGTGCCCTACTATATCAACCCACGA 2838
QY 1603 GACTCAAAACAACCTTGTGGGACCATTCCCAAAATGACAGAGCTCTACCACTTTCTTTAGCTGA 1662
Db 2839 GACTCAAAACAACCTTGTGGGACCATTCCCAAAATGACAGAGCTCTACCACTTTCTTTAGCTGA 2898
QY 1663 CCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAAGCTGCAGAA 1722
Db 2899 CCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAAGCTGCAGAA 2958
QY 1723 GGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGTCATGTGATGCCCTTGAGCCAGCACAA 1782
Db 2959 GGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGTCATGTGATGCCCTTGAGCCAGCACAA 3018
QY 1783 CCTCAAGCAAAATGACCCAGCCCATGGATATCCTGCAGATTATTAATTTGTTGACCACTAT 1842
Db 3019 CCTCAAGCAAAATGACCCAGCCCATGGATATCCTGCAGATTATTAATTTGTTGACCACTAT 3078
QY 1843 TTATGACCCGCTGGAGCAAGAGCACAAACAATTTGGTCAACGTCCCTCTCTGCGTGGATAT 1902
Db 3079 TTATGACCCGCTGGAGCAAGAGCACAAACAATTTGGTCAACGTCCCTCTCTGCGTGGATAT 3138
QY 1903 GTGCTGAACCTGGCTGCTGATGTTTATGATACGGGACGAAACAGGGAGGATCCGTTCTCT 1962
Db 3139 GTGCTGAACCTGGCTGCTGATGTTTATGATACGGGACGAAACAGGGAGGATCCGTTCTCT 3198
QY 1963 GTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATA 2022
Db 3199 GTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATA 3258
QY 2023 CCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCTCTCT 2082
Db 3259 CCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCTCTCT 3318
QY 2083 TCTGCATGATTCTATCCAA 2101
Db 3319 TCTGCATGATTCTATCCAA 3337

RESULT 13

ABX81997

ID ABX81997 standard; DNA; 5417 BP.

XX

AC ABX81997;

XX

DT 13-AUG-2002 (first entry)
XX DNA encoding mini-dystrophin protein deltaR4-R23.
DE Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
XX Duchenne's muscular dystrophy; DMD; dystrophin; ds.
KW Homo sapiens.
XX Synthetic.
OS WO200229056-A2.
XX 11-APR-2002.
PN 04-OCT-2001; 2001WO-US31126.
XX 06-OCT-2000; 2000US-238848P.
PF (UNMI) UNIV MICHIGAN.
PR Chamberlain JS, Harper SQ;
XX WPI; 2002-435334/46.
DR A composition for preparing therapeutic drugs, has a mini-dystrophin
XX peptide comprising a specific number of spectrin-like repeat domains,
XX or a nucleic acid sequence encoding the mini-dystrophin peptide -
XX Disclosure; Fig 12; 145pp; English.
XX The invention describes a composition comprising a mini-dystrophin
CC peptide comprising a spectrin-like repeat domain, where the domain
CC comprises n spectrin-like repeats, and contains no more than n
CC spectrin-like repeats, where n is an even number between 4-24, or a
CC nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin
CC peptide or the polynucleotide encoding it is useful as a medicament,
CC for preparing a drug for therapeutic application and in the preparation
CC of a composition for treatment of muscle disease, e.g. Duchenne's
CC muscular dystrophy (DMD). This sequence represents a mini-dystrophin
CC sequence of the invention.
XX Sequence 5417 BP; 1700 A; 1192 C; 1182 G; 1343 T; 0 other;
SQ
Query Match 76.7%; Score 1611.6; DB 24; Length 5417;
Best Local Similarity 85.3%; Pred. No. 0;
Matches 1872; Conservative 0; Mismatches 229; Indels 93; Gaps 3;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCCACCACCTCTGACCCCTACCGGAGCCCAT 60
DB 1099 GAGCTATGCCTACACACAGGCTGCTTATGTCCACCACCTCTGACCCCTACCGGAGCCCAT 1158
QY 61 TCCTTCACAGCATTTGGAAAGCTCCTGAAGACAAGTATTTGGCAGTTTCATTTGATGGAGAG 120
DB 1159 TCCTTCACAGCATTTGGAAAGCTCCTGAAGACAAGTATTTGGCAGTTTCATTTGATGGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTTAGAAGAAGTATTTATCGTGGCTTCTTTC 180
DB 1219 TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTTAGAAGAAGTATTTATCGTGGCTTCTTTC 1278
QY 181 TGCTGAGGACACATTTGCAAGCAAGGAGAGATTTCTAATGATGTGGAAAGTGGTGAAGA 240
DB 1279 TGCTGAGGACACATTTGCAAGCAAGGAGAGATTTCTAATGATGTGGAAAGTGGTGAAGA 1338
QY 241 CCAGTTTCATCTCATGAGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
DB 1339 CCAGTTTCATCTCATGAGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1398
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTTATCAGAAGATGAAGA 360
DB 1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTTATCAGAAGATGAAGA 1458
QY 361 AACTGAGTACAAGAGCAGATGATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420

Db 1459 AACTGAAGTACAAGACCAGATGAATCTCTTAAATTCAGATGGGAATGCCTCAGGGTAGC 1518
QY 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 480
Db 1519 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 1578
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAAGAAAGAAACAAGGAAATGGAGGAAGA 540
Db 1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAAGAAAGAAACAAGGAAATGGAGGAAGA 1638
QY 541 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAACAACATAAGGTGCTTCA 600
Db 1639 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAACAACATAAGGTGCTTCA 1698
QY 601 AGAAGATCTAGAACAAGAACAAAGTCAAGGTCAATTTCTCTCACTCACATGGTGGTAGT 660
Db 1699 AGAAGATCTAGAACAAGAACAAAGTCAAGGTCAATTTCTCTCACTCACATGGTGGTAGT 1758
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAAACAACCTTAAGGTATTGGG 720
Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAAACAACCTTAAGGTATTGGG 1818
QY 721 AGATCGATGGCAACATCTGTAGATGGACAGAAGACCGCTGGTCTCTTTTACAAGACAG 780
Db 1819 AGATCGATGGCAACATCTGTAGATGGACAGAAGACCGCTGGTCTCTTTTACAAGACAT 1878
QY 781 TTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGACGAAACTTCTGGTGTGGCTACAGCT 840
Db 1879 CCTTCTCAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTCAATGGCTTTTCA 1938
QY 841 GAAAGATGATGAATTAAGCCGGCAGGCACCTTATTTGGAGGCGACTTTTCCAGCAGTTTCA 900
Db 1939 AAAAGAAGATGCAGTGAACAAGATTACACAACTGGCTTTAAAGATCAAAATGAAATGTT 1998
QY 901 GCAGAACGATGTACATAGGGCC--TTCAAGAGGGAATTTGAAAACCTTAAGAACCTGTAATC 958
Db 1999 ATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2058
QY 959 ATGAGTACTCTTGAGACTGTACGAATATTT-----CTGACAGAGCAGCCTTTGGA 1008
Db 2059 GGGCAAACTGTATTCTCACTCAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2118
QY 1009 AGGACTAGAGAAACTCTACCAGGAGCCCGAGAGTGCCTCCTGAGGAGAGAGCCCGAGAA 1068
Db 2119 CCAGAAGACGGAAGCATGGCTGGATAACTTTTGGCGGTGTTGGGATAATTTAGTCCAAA 2178
QY 1069 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATAC----- 1110
Db 2179 ACTTGAAAAGAGTACAGCACAGATTTTCACAGGCTGTCCACCCTCAGCCATCACTAAC 2238
QY 1111 ----- 1110
Db 2239 ACAGACAACTGTAATGGAAAACAGTAACTACGGTGACCACAAGGGAACAGATCCTGGTAA 2298
QY 1111 ---TGAGTGGGAAAAATTGAACCTTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGAC 1167
Db 2299 GCATGCTCAAGAGGAATTCACCACTCCCAAAAGAGAGGCGAGATTACTGTGGA 2358
QY 1168 CCTTGAAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCA 1227
Db 2359 TCTTGAAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCA 2418
QY 1228 AGCTGAGGTGATCAAGGATTCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCA 1287
Db 2419 AGCTGAGGTGATCAAGGATTCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCA 2478
QY 1288 AGATCACCTCGAGAAAAGTCAAGGCACCTTCAGGAGAAAATTCGCCCTCTGAAAGAGAACGT 1347
Db 2479 AGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGAAAATTCGCCCTCTGAAAGAGAACGT 2538
QY 1348 GAGCCACGTCAATGACCTTGTCTGCCAGCTTACCCTTTGGGCATTTAGCTCTCACCGTA 1407
Db 2539 GAGCCACGTCAATGACCTTGTCTGCCAGCTTACCCTTTGGGCATTTAGCTCTCACCGTA 2598

QY 1408 TAACCTCAGCACTCTGGAGACCTGAAACACACAGATGGAAGCTTCTGAGGTGGCCGTCGA 1467
Db |||||
QY 1468 GGACCGAGTCAGGCAGCTGCTGAAGCCACACAGGACTTGGTCCAGCATCTCAGCACTT 1527
Db |||||
QY 1528 TCTTTCCACGTCGTCTCCAGGTCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTA 1587
Db |||||
QY 1588 CTATATCAACACGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGACTCTA 1647
Db |||||
QY 1648 CCAGTCTTTAGCTGACCTGAATATGTCTCAGCTTATAGGACTGCCATGAAACT 1707
Db |||||
QY 1708 CCGAAGACTGCAGAAAGCCCTTTGCTTGGATCTCTTGAGCCTGTCTGCTGCTGATGC 1767
Db |||||
QY 1768 CTTGGACCAACACCTCAAGCAAAATGACCGCCCATGGATATCTCTGAGATTATTA 1827
Db |||||
QY 1828 TTGTTTGACCACTATTTATGACCGCCTGGAGCAAGACACAACTTTGGTCAACGTCCT 1887
Db |||||
QY 1888 TCTCTGCGTGATATGTGTCTGAACCTGCTGCTGAATCTTTATGATACGGGACGACAGG 1947
Db |||||
QY 1948 GAGGATCCGTGCTGCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGA 2007
Db |||||
QY 2008 AGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGG 2067
Db |||||
QY 2068 CAGGCTGGGCTCCTCTGCTGATGATTTCTATCCAA 2101
Db |||||
QY 2101 CAGGCTGGGCTCCTCTGCTGATGATTTCTATCCAA 3292
Db |||||
RESULT 14
AAD37234
ID AAD37234 standard; DNA; 3999 BP.
XX
AC AAD37234;
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta3990.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
CS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.

XX (XIAO/) XIAO X.
PI Xiao X;
XX WPI; 2002-049342/06.
DR
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 46-47; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3990 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1 and R2), 8059-10227 (rods R22, R23 and R24, hinge
CC H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids).
XX
SQ Sequence 3999 BP; 1223 A; 907 C; 933 G; 936 T; 0 other;
Query Match 76.2%; Score 1602; DB 24; Length 3999;
Best Local Similarity 81.1%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 489; Gaps 1;
QY 1 GAGCTATGCTTACACACAGGCTGCTTATGTCCACCACCTCTGACCCCTACACGGAGCCCAAT 60
Db |||||
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGSCAGTTTCATTGATGGAGAG 120
Db |||||
QY 121 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAAGATTTATCGTGGCTTCTTTC 180
Db |||||
QY 181 TGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db |||||
QY 241 CCAGTTTTCATCTACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCCCGGTTGG 300
Db |||||
QY 301 TAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
Db |||||
QY 361 AACTGAAGTACAAAGACAGATGAATCTCCTAAATTCAGATGGAATGCCCTCAGGGTAGC 420
Db |||||
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGATTTTAAATGGATCTCCAGAAATCAGAAACT 480
Db |||||
QY 481 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAACAGGAAAATGGAGGAAGA 540
Db |||||
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCCAAGTACAACATAAGGTGCTTCA 600
Db |||||
QY 600 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCCAAGTACAACATAAGGTGCTTCA 1499
Db |||||

QY 601 AGAAGATCTAGAAACAAGAACAAAGTCAAGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 660
DB 1500 AGAAGATCTAGAAACAAGAACAAAGTCAAGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 1559
QY 661 TGATGAATCTAGTGGAGATCACCGAACTGCTGCTTTGGAAAGAACAACTTAAAGGTATTGGG 720
DB 1560 TGATGAATCTAGTGGAGATCACCGAACTGCTGCTTTGGAAAGAACAACTTAAAGGTATTGGG 1619
QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGA--- 777
DB 1620 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACCA 1679
QY 778 ----- 777
DB 1680 GCCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTCTCCTACTCAGACTGTTACTCT 1739
QY 778 ----- 777
DB 1740 GGTGACACAACTGTGGTTACTAAGGAAACTGCCATCTCCAAACTAGAAAATGCCATCTTC 1799
QY 778 ----- 777
DB 1800 CTTGATGTTGGAGTACCTACTCATAGATTACTGTCAACAGTTCCTCCCTGGACCTGGAAAA 1859
QY 778 ----- 777
DB 1860 GTTTCTTGCCCTGGCTTACAGAAGCTGAAACAACACTGCCAATGTCTACAGGATGCTACCCG 1919
QY 778 ----- 777
DB 1920 TAAGGAAGGCTCCTAGAACTCCAAAGGGAGTAAAGAGCTGATGAAACAATGSCAAGA 1979
QY 778 ----- 777
DB 1980 CCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAACAGCCA 2039
QY 778 ----- 777
DB 2040 AAAAATCCTGAGATCCCTGGAAGGTTCGGATGATGCAGTCTCTGTTACAAAGACGTTTGA 2099
QY 778 ----- 777
DB 2100 TAACATGAATCAAGTGGAGTGAATTCGGAATAAGTCTCTCAACATTAGTCCCATTT 2159
QY 778 -----CAGTTCTGACCAAGTGAAGCGCTGTGCACCTTTCTCTGCAGGAACCTTCTGGTGTG 831
DB 2160 GGAAGCCAGTCTGACCAAGTGAAGCGCTGTGCACCTTTCTCTGCAGGAACCTTCTGGTGTG 2219
QY 832 GCTACAGCTGAAAGATGATGAATTAAAGCCGGCAGGCACCTATTTGGAGGCGACTTTCCAGC 891
DB 2220 GCTACAGCTGAAAGATGATGAATTAAAGCCGGCAGGCACCTATTTGGAGGCGACTTTCCAGC 2279
QY 892 AGTTCAGAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACTAAAGAAC 951
DB 2280 AGTTCAGAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACTAAAGAAC 2339
QY 952 TGTAATCATGACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGG 1011
DB 2340 TGTAATCATGACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGG 2399
QY 1012 ACTAGAGAACTCTACCAAGGAGCCCGAGAGCTGCCTCCTGAGAGAGAGCCCGAGATGT 1071
DB 2400 ACTAGAGAACTCTACCAAGGAGCCCGAGAGCTGCCTCCTGAGAGAGAGCCCGAGATGT 2459
QY 1072 CACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCT 1131
DB 2460 CACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCT 2519
QY 1132 GCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTCA 1191
DB 2520 GCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTCA 2579

QY 1192 AGAGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTG 1251
DB 2580 AGAGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTG 2639
QY 1252 GCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACTCGAGAAAAGTCAAGGC 1311
DB 2640 GCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACTCGAGAAAAGTCAAGGC 2699
QY 1312 ACTTCGAGGAGAAAATTGCGCTCTGAAAGAGAACGTCGAGCCACGTCATGACCTTGCTCG 1371
DB 2700 ACTTCGAGGAGAAAATTGCGCTCTGAAAGAGAACGTCGAGCCACGTCATGACCTTGCTCG 2759
QY 1372 CCAGCTTACCACCTTTGGGCAATTCAGCTCTCAACGTAFAAACCTCAGCACTCTGGAAGACCT 1431
DB 2760 CCAGCTTACCACCTTTGGGCAATTCAGCTCTCAACGTAFAAACCTCAGCACTCTGGAAGACCT 2819
QY 1432 GAACACCAAGATGGAAGCTTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGA 1491
DB 2820 GAACACCAAGATGGAAGCTTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGA 2879
QY 1492 AGCCCAAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCC 1551
DB 2880 AGCCCAAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCC 2939
QY 1552 CTGGGAGAGAGCCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAAC 1611
DB 2940 CTGGGAGAGAGCCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAAC 2999
QY 1612 AACITGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTTAGCTGACCTGAATAA 1671
DB 3000 AACITGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTTAGCTGACCTGAATAA 3059
QY 1672 TGTCAATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTG 1731
DB 3060 TGTCAATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTG 3119
QY 1732 CTTGGATCTCTTGAGCCTGTCAAGTGCATGTGATGCCCTTGACCAGCAACACCTCAAGCA 1791
DB 3120 CTTGGATCTCTTGAGCCTGTCAAGTGCATGTGATGCCCTTGACCAGCAACACCTCAAGCA 3179
QY 1792 AAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTATGACCG 1851
DB 3180 AAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTATGACCG 3239
QY 1852 CCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTCGGTGGATATGTGTCTGAA 1911
DB 3240 CCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTCGGTGGATATGTGTCTGAA 3299
QY 1912 CTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCTCTTTTAA 1971
DB 3300 CTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCTCTTTTAA 3359
QY 1972 AACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAA 2031
DB 3360 AACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAA 3419
QY 2032 GCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGCAGGCTGGGCTCCTTCTGCAATGA 2091
DB 3420 GCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGCAGGCTGGGCTCCTTCTGCAATGA 3479
QY 2092 TTCTATCCAA 2101
DB 3480 TTCTATCCAA 3489

RESULT 15
AAD37256
ID AAD37256 standard; DNA; 4966 BP.
XX
AC AAD37256;
XX
DT 21-AUG-2002 (first entry)
XX

DE Adeno-associated virus vector plasmid, AAV-MCK-delta3990.
XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Unidentified.
XX
XX WO200183695-A2.
XX
PD 08-NOV-2001.
XX
XX 27-APR-2001; 2001WO-US13677.
PF
XX
PR 28-APR-2000; 2000US-200777P.
XX
XX (XIAO/) XIAO X.
PA
XX
PI Xiao X;
XX
XX WPI; 2002-049342/06.
DR
XX
XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
XX Example 1; Page 59-60; 71pp; English.
PS
XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence.
XX
SQ Sequence 4966 BP; 1403 A; 1216 C; 1232 G; 1115 T; 0 other;

Query Match 76.2%; Score 1602; DB 24; Length 4966;
Best Local Similarity 81.1%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 489; Gaps 1;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTACACCACTCTGACCCCTACACGGAGCCCAT 60
DB 1657 GAGCTATGCCTACACACAGGCTGCTTATGTACACCACTCTGACCCCTACACGGAGCCCAT 1716
QY 61 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCATTGATGGAGAG 120
DB 1717 TCCTTCACAGCATTTTGAAGCTCCTGAAGACAAAGTCAATTTGGCAGTTCATTGATGGAGAG 1776
QY 121 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
DB 1777 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1836
QY 181 TGCTGAGGACACATTGCAACCAACAGGAGAGATTTCATATGATGTGGAAGTGGTGAAGA 240
DB 1837 TGCTGAGGACACATTGCAACCAACAGGAGAGATTTCATATGATGTGGAAGTGGTGAAGA 1896
QY 241 CCAGTTTCATCTACTGAGGGGTACATGATGGATTTCACAGCCCATCAGGCCCGGTTGG 300
DB 1897 CCAGTTTCATCTACTGAGGGGTACATGATGGATTTCACAGCCCATCAGGCCCGGTTGG 1956
QY 301 TAATATTCTACAAATTTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 360
DB 1957 TAATATTCTACAAATTTGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 2016

QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 420
DB 2017 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCTCAGGGTAGC 2076
QY 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAAACT 480
DB 2077 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAAACT 2136
QY 481 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAAAGAAATGGAGGAAGA 540
DB 2137 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAAAGAAATGGAGGAAGA 2196
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACAACATAAGGTGCTTCA 600
DB 2197 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACAACATAAGGTGCTTCA 2256
QY 601 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGTAGT 660
DB 2257 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGTGTAGT 2316
QY 661 TGATGAATCTAGTGGAGATCACGCCAACTGCTGCTTTGGAAGAAACAACCTTAAGGTATTGGG 720
DB 2317 TGATGAATCTAGTGGAGATCACGCCAACTGCTGCTTTGGAAGAAACAACCTTAAGGTATTGGG 2376
QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGCCGCTGGTCTTTTACAAGA--- 777
DB 2377 AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGCCGCTGGTCTTTTACAAGACCA 2436
QY 778 ----- 777
DB 2437 GCCTGACCTAGCTCCTGGACTGACCACCTATTGGAGCCTCTCTACTCAGACTGTACTCT 2496
QY 778 ----- 777
DB 2497 GGTGACACAACCTGTGTTACTAAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTTC 2556
QY 778 ----- 777
DB 2557 CTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAA 2616
QY 778 ----- 777
DB 2617 GTTTCTTGCCTGGCTTACAGAAGCTGAAACAACCTGCCAATGTCTCAGAGATGCTACCCG 2676
QY 778 ----- 777
DB 2677 TAAGGAAAGGCTCCTAGAGACTCCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGA 2736
QY 778 ----- 777
DB 2737 CCTCCAAGGTGAAATGAAGCTCACACAGATGTTTATCACAACTGGATGAAACAGCCA 2796
QY 778 ----- 777
DB 2797 AAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTACAAGACGTTTGA 2856
QY 778 ----- 777
DB 2857 TAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCCATT 2916
QY 778 -----CAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTCTGACGAACTTCTGGTGTG 831
DB 2917 GGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTCTGACGAACTTCTGGTGTG 2976
QY 832 GCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGC 891
DB 2977 GCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGC 3036
QY 892 AGTTTCAGAAGCAGAACCGATGTACATAGGGCCTTCAAGAGGGAATTTGAAACTAAAGAACC 951
DB 3037 AGTTTCAGAAGCAGAACCGATGTACATAGGGCCTTCAAGAGGGAATTTGAAACTAAAGAACC 3096
QY 952 TGTAAATCATGAGTACTCTTTGAGACTGTACGAAATTTCTGACAGAGCAGCCTTTGGAAGG 1011

Db 3097 TGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCTTTGGAAAGG 3156
QY 1012 ACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCTCCCTGAGGAGAGAGCCCAAGATGT 1071
Db 3157 ACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCTCCCTGAGGAGAGAGCCCAAGATGT 3216
QY 1072 CACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCT 1131
Db 3217 CACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCT 3276
QY 1132 GCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCA 1191
Db 3277 GCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCA 3336
QY 1192 AGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTG 1251
Db 3337 AGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTG 3396
QY 1252 GCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGC 1311
Db 3397 GCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGC 3456
QY 1312 ACTTCGAGGAGAAATTGGCGCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCG 1371
Db 3457 ACTTCGAGGAGAAATTGGCGCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCG 3516
QY 1372 CCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCT 1431
Db 3517 CCAGCTTACCACCTTTGGGCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCT 3576
QY 1432 GAACACAGATGGAAGCTTCTGAGGTGGCCGTGAGAGCCGAGTCAGGCAGCTGCATGA 1491
Db 3577 GAACACAGATGGAAGCTTCTGAGGTGGCCGTGAGAGCCGAGTCAGGCAGCTGCATGA 3636
QY 1492 AGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCC 1551
Db 3637 AGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCC 3696
QY 1552 CTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAC 1611
Db 3697 CTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAC 3756
QY 1612 AACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAA 1671
Db 3757 AACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAA 3816
QY 1672 TGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAGGCCCTTTG 1731
Db 3817 TGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAGGCCCTTTG 3876
QY 1732 CTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCCCTTGGACCGACACACCTCAAGCA 1791
Db 3877 CTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCCCTTGGACCGACACACCTCAAGCA 3936
QY 1792 AAATGACCAAGCCCATGGATATCCTGCAAGATTATTAATTGTTTGACCACCTATTATGACCG 1851
Db 3937 AAATGACCAAGCCCATGGATATCCTGCAAGATTATTAATTGTTTGACCACCTATTATGACCG 3996
QY 1852 CCTGGAGCAAGAGCACAAACATTTGGTCAACGTCCCTCTCTGCGTGGATATGTGTCTGAA 1911
Db 3997 CCTGGAGCAAGAGCACAAACATTTGGTCAACGTCCCTCTCTGCGTGGATATGTGTCTGAA 4056
QY 1912 CTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGTCTTTTAA 1971
Db 4057 CTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGTCTTTTAA 4116
QY 1972 AACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGTACCTTTTCAA 2031
Db 4117 AACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGTACCTTTTCAA 4176
QY 2032 GCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGAGGCTGGGCCCTCCTTCTGCATGA 2091

Db 4177 GCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGAGCTGGGCCCTCCTTCTGCATGA 4236
QY 2092 TTCTATCCAA 2101
Db 4237 TTCTATCCAA 4246

Search completed: February 1, 2004, 12:05:04
Job time : 530.045 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 11:11:40 ; Search time 129.879 Seconds
(without alignments)
7140.092 Million cell updates/sec

Title: US-09-845-416-12_COPY_900_3000
Perfect score: 2101
Sequence: 1 gagctatgcctacacacagg.....ttctgcatgattctatccaa 2101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|---------|-------|---------------------|
| 1 | 1328.8 | 63.2 | 5952 | 4 | US-09-687-875A-1 |
| 2 | 1317.8 | 62.7 | 13977 | 4 | US-09-484-970B-60 |
| C 3 | 1149.6 | 54.7 | 19307 | 3 | US-08-836-022A-10 |
| C 4 | 1149.6 | 54.7 | 19307 | 3 | US-09-427-048A-10 |
| 5 | 547.4 | 26.1 | 6045 | 4 | US-09-091-501B-7 |
| 6 | 547.4 | 26.1 | 10320 | 4 | US-09-091-501B-9 |
| 7 | 79.4 | 3.8 | 200 | 4 | US-09-091-501B-5 |
| 8 | 78.6 | 3.7 | 200 | 4 | US-09-091-501B-4 |
| 9 | 78.6 | 3.7 | 200 | 4 | US-09-091-501B-6 |
| C 10 | 76.6 | 3.6 | 7218 | 1 | US-08-232-463-14 |
| 11 | 44.2 | 2.1 | 2574 | 4 | US-09-668-313A-10 |
| C 12 | 44 | 2.1 | 1230025 | 4 | US-09-198-452A-1 |
| 13 | 43.4 | 2.1 | 1179 | 4 | US-09-107-532A-1186 |
| 14 | 42.8 | 2.0 | 1690 | 4 | US-09-620-312D-69 |
| 15 | 42.8 | 2.0 | 7812 | 3 | US-09-368-590-1 |
| 16 | 40.4 | 1.9 | 2223 | 1 | US-08-257-073-4 |
| 17 | 39.2 | 1.9 | 16995 | 4 | US-08-961-527-82 |
| 18 | 38.6 | 1.8 | 1995 | 1 | US-08-425-069-3 |
| 19 | 38.6 | 1.8 | 1995 | 2 | US-08-317-844B-3 |
| 20 | 38.4 | 1.8 | 7672 | 4 | US-09-220-132-24 |
| 21 | 38.2 | 1.8 | 428 | 4 | US-09-668-313A-3 |
| 22 | 38.2 | 1.8 | 1131 | 6 | 5180810-3 |
| 23 | 38.2 | 1.8 | 1784 | 6 | 5180810-2 |
| 24 | 38.2 | 1.8 | 4439 | 4 | US-09-668-313A-17 |
| C 25 | 38 | 1.8 | 1394 | 4 | US-09-247-155-76 |
| 26 | 36.8 | 1.8 | 1886 | 6 | 5210183-1 |
| 27 | 36.6 | 1.7 | 1845 | 4 | US-08-887-534A-22 |

| | | | | | | |
|----|------|-----|--------|---|--------------------|-------------------|
| 28 | 36.6 | 1.7 | 1845 | 4 | US-09-527-431-22 | Sequence 22, Appl |
| 29 | 36.2 | 1.7 | 289 | 3 | US-09-007-005-17 | Sequence 17, Appl |
| 30 | 36.2 | 1.7 | 289 | 3 | US-09-244-796-17 | Sequence 17, Appl |
| 31 | 36.2 | 1.7 | 2447 | 2 | US-09-014-969-14 | Sequence 14, Appl |
| 32 | 36.2 | 1.7 | 168575 | 4 | US-09-426-290-1 | Sequence 1, Appl |
| 33 | 36 | 1.7 | 2873 | 4 | US-08-630-915A-193 | Sequence 193, App |
| 34 | 35.8 | 1.7 | 790 | 3 | US-09-461-474-13 | Sequence 13, Appl |
| 35 | 35.6 | 1.7 | 4868 | 1 | US-08-139-937-12 | Sequence 12, Appl |
| 36 | 35.6 | 1.7 | 4868 | 5 | PCT-US93-11310-12 | Sequence 12, Appl |
| 37 | 35.6 | 1.7 | 8257 | 4 | US-09-595-684B-30 | Sequence 30, Appl |
| 38 | 35.6 | 1.7 | 8789 | 1 | US-08-328-254-5 | Sequence 5, Appl |
| 39 | 35.6 | 1.7 | 10136 | 1 | US-08-353-700-2 | Sequence 2, Appl |
| 40 | 35.6 | 1.7 | 10136 | 5 | PCT-US95-16216-2 | Sequence 2, Appl |
| 41 | 35.2 | 1.7 | 3466 | 1 | US-08-468-036-38 | Sequence 38, Appl |
| 42 | 35.2 | 1.7 | 3466 | 2 | US-08-376-843-38 | Sequence 38, Appl |
| 43 | 35 | 1.7 | 4766 | 5 | PCT-US93-07261-10 | Sequence 10, Appl |
| 44 | 34.8 | 1.7 | 714 | 4 | US-09-328-352-4055 | Sequence 4055, Ap |
| 45 | 34.8 | 1.7 | 1453 | 4 | US-08-714-741-33 | Sequence 33, Appl |

ALIGNMENTS

RESULT 1
US-09-687-875A-1
; Sequence 1, Application US/09687875A
; Patent No. 6544786
; GENERAL INFORMATION:
; APPLICANT: Xiao, Xiao
; APPLICANT: Liu, Paul
; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPLICED PE
; FILE REFERENCE: 00792
; CURRENT APPLICATION NUMBER: US/09/687,875A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/158,868
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 5952
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2897)..(2898)
; OTHER INFORMATION: S4 junction site
; NAME/KEY: misc feature
; LOCATION: (3198)..(3199)
; OTHER INFORMATION: S2 junction site
US-09-687-875A-1

Query Match 63.2%; Score 1328.8; DB 4; Length 5952;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

| | | | |
|----|------|--|------|
| QY | 762 | GGTTCTTTTACAAGACAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAAC | 821 |
| Db | 3284 | GGTCCATTGGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAAC | 3343 |
| QY | 822 | TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCG | 881 |
| Db | 3344 | TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCG | 3403 |
| QY | 882 | ACTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAA | 941 |
| Db | 3404 | ACTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAA | 3463 |
| QY | 942 | CTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTCTGCAGAGCAGC | 1001 |
| Db | 3464 | CTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTCTGCAGAGCAGC | 3523 |
| QY | 1002 | CTTTGGAAGGACTAGAGAACTCTACCAGGAGGCCAGAGAGCTGCCTCTCTGAGGAGAG | 1061 |

Db 3524 CTTTGAAGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG 3583

Qy 1062 CCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGAAA 1121

Db 3584 CCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGAAA 3643

Qy 1122 AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAAGACTCC 1181

Db 3644 AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAAGACTCC 3703

Qy 1182 AGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCCCAAGCTCAGGTGATCA 1241

Db 3704 AGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCCCAAGCTCAGGTGATCA 3763

Qy 1242 AGGATCCTGGCAGCCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGA 1301

Db 3764 AGGATCCTGGCAGCCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGA 3823

Qy 1302 AAGTCAAGGCACCTTCGAGGAGAAAATTGGCGCTCTGAAAGAGAAAGTGAAGCCACGTCAATG 1361

Db 3824 AAGTCAAGGCACCTTCGAGGAGAAAATTGGCGCTCTGAAAGAGAAAGTGAAGCCACGTCAATG 3883

Qy 1362 ACCTTGCTCGCCAGCTTACACCTTTGGGCAATTCAGCTCTCACCGTATAAAGCTCAGCACTC 1421

Db 3884 ACCTTGCTCGCCAGCTTACACCTTTGGGCAATTCAGCTCTCACCGTATAAAGCTCAGCACTC 3943

Qy 1422 TGGAAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAGGC 1481

Db 3944 TGGAAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAGGC 4003

Qy 1482 AGCTGCATGAAGCCCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTG 1541

Db 4004 AGCTGCATGAAGCCCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTG 4063

Qy 1542 TCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACG 1601

Db 4064 TCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACG 4123

Qy 1602 AGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTG 1661

Db 4124 AGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTG 4183

Qy 1662 ACCTGAATAATGTGATCTCAGCTTATAGGACTGGCCATGAAACTCCGAAGACTGCAGA 1721

Db 4184 ACCTGAATAATGTGATCTCAGCTTATAGGACTGGCCATGAAACTCCGAAGACTGCAGA 4243

Qy 1722 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGACCCAGCACA 1781

Db 4244 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGACCCAGCACA 4303

Qy 1782 ACCTCAAGCAAAATGACCGCCCATGGATATCCTGACAGATTATTAATTGTTTGACCACTA 1841

Db 4304 ACCTCAAGCAAAATGACCGCCCATGGATATCCTGACAGATTATTAATTGTTTGACCACTA 4363

Qy 1842 TTTATGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCCTCTCTGCGTGGATA 1901

Db 4364 TTTATGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCCTCTCTGCGTGGATA 4423

Qy 1902 TGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCC 1961

Db 4424 TGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCC 4483

Qy 1962 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAGAT 2021

Db 4484 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAGAT 4543

Qy 2022 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGGAGGCTGGGCTTCC 2081

Db 4544 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGGAGGCTGGGCTTCC 4603

Qy 2082 TTCTGCATGATTTCTATCCAA 2101

Db 4604 TTCTGCATGATTTCTATCCAA 4623

RESULT 2

US-09-484-970B-60

; Sequence 60, Application US/09484970B

; Patent No. 6426186

; GENERAL INFORMATION:

; APPLICANT: Jones, Karen A.

; APPLICANT: Volkmuth, Wayne

; APPLICANT: Walker, Michael G.

; TITLE OF INVENTION: BONE REMODELING GENES

; FILE REFERENCE: PB-0014 US

; CURRENT APPLICATION NUMBER: US/09/484,970B

; CURRENT FILING DATE: 2000-01-18

; NUMBER OF SEQ ID NOS: 172

; SOFTWARE: PERL Program

; SEQ ID NO 60

; LENGTH: 13977

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. 6426186 229357.11CB1

; NAME/KEY: unsure

; LOCATION: 11721-11761, 12294, 13969

; OTHER INFORMATION: a, t, c, g, or other

US-09-484-970B-60

Query Match 62.7%; Score 1317.8; DB 4; Length 13977;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 1333; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

Qy 762 GGGTCTTTTACAAGACAGTTCTGACCAGTGGAGCGTCTGCACCTTTCTCTCAGGAAC 821

Db 8598 GGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAGCGTCTGCACCTTTCTCTCAGGAAC 8657

Qy 822 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGCACCTATTGGAGCG 881

Db 8658 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGCACCTATTGGAGCG 8717

Qy 882 ACTTTCAGCAGTTTCAAGACGAAACGATGATACATAGGGCTTCAAGAGGGAATTGAAA 941

Db 8718 ACTTTCAGCAGTTTCAAGACGAAACGATGATACATAGGGCTTCAAGAGGGAATTGAAA 8777

Qy 942 CTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 1001

Db 8778 CTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 8837

Qy 1002 CTTTGAAGGACTAGAGAAAACCTTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG 1061

Db 8838 CTTTGAAGGACTAGAGAAAACCTTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAG 8897

Qy 1062 CCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAA 1121

Db 8898 CCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAA 8957

Qy 1122 AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAAGACTCC 1181

Db 8958 AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAAGACTCC 9017

Qy 1182 AGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCA 1241

Db 9018 AGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCA 9077

Qy 1242 AGGATCCTGGCAGCCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGA 1301

Db 9078 AGGATCCTGGCAGCCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGA 9137

Qy 1302 AAGTCAAGGCACCTTCGAGGAGAAAATTGGCCCTCTGAAAGAGAACGCTGAGCCACGTCAATG 1361

Db 9138 AAGTCAAGGCACCTTCGAGGAGAAAATTGGCCCTCTGAAAGAGAACGCTGAGCCACGTCAATG 9197

Qy 1362 ACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTTCAGCTCTCACCGTATAACCTCAGCACTC 1421

Db 9198 ACCTTGCTCGCAGCTTACCACITTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTC 9257
QY 1422 TGGAGAGACCTGAAACACACAGATGGAAGCTTCTGCAGGTGGCGCTCGAGGACCGAGTCAGGC 1481
Db 9258 TGGAGAGACCTGAAACACACAGATGGAAGCTTCTGCAGGTGGCGCTCGAGGACCGAGTCAGGC 9317
QY 1482 AGCTGCATGAAGCCCCACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTG 1541
Db 9318 AGCTGCATGAAGCCCCACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTG 9377
QY 1542 TCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCCCTACTATATATCAACACG 1601
Db 9378 TCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACACG 9437
QY 1602 AGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTTACCAGTCTTTAGCTG 1661
Db 9438 AGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTTACCAGTCTTTAGCTG 9497
QY 1662 ACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 1721
Db 9498 ACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 9557
QY 1722 AGGCCCTTTGCTTGATCTCTTGAGCCTGTTCAGCTGTCATGTGATGCCCTTGGACCAACACA 1781
Db 9558 AGGCCCTTTGCTTGATCTCTTGAGCCTGTTCAGCTGTCATGTGATGCCCTTGGACCAACACA 9617
QY 1782 ACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTA 1841
Db 9618 ACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTA 9677
QY 1842 TTTATGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTCGGTGGATA 1901
Db 9678 TTTATGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTCGGTGGATA 9737
QY 1902 TGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACCAACAGGGAGGATCCGTGTCC 1961
Db 9738 TGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACCAACAGGGAGGATCCGTGTCC 9797
QY 1962 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACAATTTGGAAGACAAAGTACAGAT 2021
Db 9798 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACAATTTGGAAGACAAAGTACAGAT 9857
QY 2022 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTACACGCGCAGGCT-GGGCCTC 2080
Db 9858 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTACACGCGCAGGCTGGGGCCTC 9917
QY 2081 CTCTGCAATGATCTATCCAA 2101
Db 9918 CTCTGCAATGATCTATCCAA 9938

RESULT 3
US-08-836-022A-10/c
; Sequence 10, Application US/08836022A
; Patent No. 6001557
; GENERAL INFORMATION:
; APPLICANT: Trustees of the University of Pennsylvania
; APPLICANT: Wilson, James M.
; APPLICANT: Fisher, Krishna J.
; APPLICANT: Chen, Shu-Jen
; APPLICANT: Weitzman, Matthew
; TITLE OF INVENTION: Improved Adenovirus Virus and
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P O Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,022A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/331,381
; FILING DATE: 28-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GNVEN.008PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19307 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
US-08-836-022A-10

Query Match 54.7%; Score 1149.6; DB 3; Length 19307;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 1221; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY 762 GGGTTCTTTTACAAGACAGTTCTGACCAGTGAAGCGTCTGCACCTTTTCTCTGCAGGAAC 821
Db 6096 GGTCCATTGGAAGCAAGTTCTGACCAGTGAAGCGTTTGCATCTTTCTCTTCAGGAAC 6037
QY 822 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCG 881
Db 6036 TTCTTGTGTGGCTACAGCTGAAAGATGATGAATGAGCCCTCAGGCACCCATCGGTGTG 5977
QY 882 ACTTTCAGCAGTTTCAAGAGCAGACGATGTACATAGGSCCTTCAAGAGGGAATTGAAAA 941
Db 5976 ATTTCCAGCAGTTTCAAGAGCAGACGATGATATACATAGGSCCTTCAAGAGGGAATTGAAAA 5917
QY 942 CTAAGAAACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGC 1001
Db 5916 CTAAGAAACCTGTAATCATGAGTACTCTTTGAGACTGTGAGAAATATTTCTGACAGAGCAGC 5857
QY 1002 CTTTGAAGGACTAGAGAAACTCTACAGGAGCCACAGAGCTGCCTCTCTGAGGAGAGAG 1061
Db 5856 CTTTGAAGGACTAGAGAAACTCTACAGGAGCCACAGAGAACTGCCTCTCTGAGAGAGAG 5797
QY 1062 CCCAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAA 1121
Db 5796 CTCAGAATGTCACTCGGCTCCTACGAAAGCAGGCTGAGGAGGTCAACGCTGAATGGGACA 5737
QY 1122 AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCC 1181
Db 5736 AATTGAACCTGCCTCAGCTGATTGGCAGAGAAAAATAGATGAAGCTCTTGAAGACTCC 5677
QY 1182 AGGAACCTCAAGAGGCGCACCGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCA 1241
Db 5676 AGGAACCTCAGGAAGCTGCCGATGAACCTGGACCTCAAGTTGCGCCAAGCTGAGGTGATCA 5617
QY 1242 AGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGA 1301
Db 5616 AGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTGCAAGATCACCTTGAAA 5557
QY 1302 AAGTCAAGGCACCTTCGAGGAGAAATTTGGCCCTCTGAGAAAGAGAACGTGAGCCACGTCATG 1361
Db 5556 AAGTCAAGGCACCTTCGCGGAGAAATTTGCACCTCTTAAAGAGAAATGTCAATCGTGTCAATG 5497
QY 1362 ACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTGAGCTCTCACCGTATAACCTCAGCACTC 1421
Db 5496 ACCTTGACATCAGCTGACCACTGGGCATTTGAGCTCTCACCTTATAACCTCAGCACTT 5437

QY 1422 TGAAGACCTGAACACACAGATGGAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAGGC 1481
Db 5436 TGAAGATCTGAATACAGATGGAGGCTTCTACAGGTGGCTGTGGAGGACCGGTGTAGAC 5377
QY 1482 AGCTGCATGAAGCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTGTG 1541
Db 5376 AGCTGCATGAAGCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACTTCAG 5317
QY 1542 TCCAGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACAG 1601
Db 5316 TTCAGGTCCCTGGGAGAGAGCCATCTCACCACAAACAAAGTGCCCTACTATATCAACACAG 5257
QY 1602 AGACTCAAAACAACTTGCTGGGACCATCCCAAATGACAGAGCTCTACCACTCTTTAGCTG 1661
Db 5256 AGACCCAAACCACTTTGTTGGGACCAACCCCAAATGACAGAGCTCTACCACTCTTTAGCTG 5197
QY 1662 ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAGACTGCAGA 1721
Db 5196 ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAGACTGCAGA 5137
QY 1722 AGGCCCTTTGCTTGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGGACGACACA 1781
Db 5136 AGGCCCTTTGCTTGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGGACGACACA 5077
QY 1782 ACCTCAAGCAAAATGACCGCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTA 1841
Db 5076 ACCTCAAGCAAAATGACCGCCATGGATATCCTGCAGATAATTAACCTGTTGACTACAA 5017
QY 1842 TTTATGACCGCCTGGAGCAAGAGACACAACAATTTGGTCAACGTCCCTCTCTGCGTGGATA 1901
Db 5016 TTTATGATCGTCTGGAGCAAGAGACACAACAATCTGGTCAATGTCCCTCTCTGTTGGATA 4957
QY 1902 TGTGCTGAAGTGGCTGCTGATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCC 1961
Db 4956 TGTGCTCAAGTGGCTTCTCAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCC 4897
QY 1962 TGTCTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGAAGACAAGTACAGAT 2021
Db 4896 TGTCTTTAAACTGGCATCATTTCTCTGTGTAAAGCACATTTTGAAGACAAGTACAGAT 4837
QY 2022 ACCTTTTCAAGCAAGTGGCAAGTCAACAGGATTTTGTGACCGAGCGCTGGGCCTCC 2081
Db 4836 ACCTTTTCAAGCAAGTGGCAAGTCAACTGGCTTTTGTGACCGAGCGTAGGCTGGGTCTTC 4777
QY 2082 TTCTGCATGATTTCTATCCAA 2101
Db 4776 TTCTGCATGATTTCTATTCAA 4757

RESULT 4
US-09-427-048A-10/C
; Sequence 10, Application US/09427048A
; Patent No. 6203975
; GENERAL INFORMATION:
; APPLICANT: Trustees of the University of Pennsylvania
; Fisher, Krishna J.
; Chen, Shu-Jen
; Weitzman, Matthew
; TITLE OF INVENTION: Improved Adenovirus Virus and
; Methods of Use Thereof
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P O Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/427,048A
FILING DATE: 21-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/836,022
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNVN.008PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-427-048A-10

Query Match 54.7%; Score 1149.6; DB 3; Length 19307;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 1221; Conservative 0; Mismatches 119; Indels 0; Gaps 0;
QY 762 GGGTCTTTTACAAGACAGTCTTGACCAGTGGAGCGTCTGCACCTTTCTCTCAGGAAC 821
Db 6096 GGTCCTATTGGAAGCAAGTCTGACCAGTGGAGCGTTTGATCTTTCTCTCAGGAAC 6037
QY 822 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTTATGGAGCG 881
Db 6036 TTCTGTGTTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTTATGGTGTG 5977
QY 882 ACTTCCAGCAGTTTCAAGACGAGAACGATGTATAGGGCCTTCAAGAGGGGAATTGAAA 941
Db 5976 ATTTCCAGCAGTTTCAAGACGAGAACGATGTATAGGGCCTTCAAGAGGGGAATTGAAA 5917
QY 942 CTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGCACAGCAGC 1001
Db 5916 CTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTGAGAAATTTCTGCACAGCAGC 5857
QY 1002 CTTTGGAGGACTAGAGAACTCTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAG 1061
Db 5856 CTTTGGAGGACTAGAGAACTCTACCAGGAGCCAGAGAACTGCCTCCTGAGAAAGAGAG 5797
QY 1062 CCCAGAAATGTCACTCGGCTTTACGAAAGCAGGCTGAGGAGGTCAATATCTAGTGGGAAA 1121
Db 5796 CTCAGAAATGTCACTCGGCTTTACGAAAGCAGGCTGAGGAGGTCAACGCTGAATGGGACA 5737
QY 1122 AATTGAACCTGCACCTCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGACTCC 1181
Db 5736 AATTGAACCTGCCTCAGCTGATTGGCAGAGAAAATAGATGAAGCTCTTGAAGACTCC 5677
QY 1182 AGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCCAAAGCTGAGGTGATCA 1241
Db 5676 AGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCCAAAGCTGAGGTGATCA 5617
QY 1242 AGGATCTTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGA 1301
Db 5616 AGGATCTTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGA 5557
QY 1302 AAGTCAAGGCACTTCGAGGAGAAAATTGCGCCTCTGAAAGAGAACGTCGAGCCACGTCATG 1361
Db 5556 AAGTCAAGGCACTTCGAGGAGAAAATTGCGCCTCTTAAAGAGAAATGTCAATCGTGTCAATG 5497
QY 1362 ACCTTGTCTGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATACCTCAGCACTC 1421
Db 5496 ACCTTGCACATCAGCTGACCACACTGGGCAATTCAGCTCTCACCTTATACCTCAGCACTT 5437

QY 1422 TGAAGAGCCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGC 1481
Db |||||
QY 5436 TGAAGATCTGAATACCAAGATGGAGGCTTCTACAGGTGGCTGTGGAGGACCGGTGCAGAC 5377
Db |||||
QY 1482 AGCTGCATGAAGCCACAGGAGCTTGGTCCAGCATCTCAGCACTTCTTTCCACGCTCTG 1541
Db |||||
QY 5376 AGCTGCATGAAGCCACAGGAGCTTGGTCCAGCATCTCAGCACTTCTTTCCACCTCAG 5317
Db |||||
QY 1542 TCCAGGTCCTGGGAGAGAGCCATCTCGCCAAACAAAGTCCCTACTATATCAACACG 1601
Db |||||
QY 5316 TTCAGGTCCTGGGAGAGAGCCATCTCACCACAAACAAAGTCCCTACTATATCAACACG 5257
Db |||||
QY 1602 AGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTG 1661
Db |||||
QY 5256 AGACCAAAACCACTTGTGGGACCAACCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTG 5197
Db |||||
QY 1662 ACCTGAATAATGTCTAGCTTCTCAGCTTATAGGAGTCCCATGAAACTCCGAAGACTGCAGA 1721
Db |||||
QY 5196 ACCTGAATAATGTCTAGCTTCTCAGGTTCTCCGGTATAGGACTGCCATGAAGCTCAGAAGGCTCCAGA 5137
Db |||||
QY 1722 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTCTAGCTGTGATGCTTGGACCAAGCACA 1781
Db |||||
QY 5136 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTCTAGCTGTGATGCTTGGACCAAGCACA 5077
Db |||||
QY 1782 ACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGAGATTAATTAATGTTTGAACCACTA 1841
Db |||||
QY 5076 ACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGAGATAAATTAATGTTTGAACCACTA 5017
Db |||||
QY 1842 TTTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCCTCTCTCGTGGATA 1901
Db |||||
QY 5016 TTTATGATCGTCTGGAGCAAGAGCAACAATCTGGTCAATGTCCCTCTCTGTGGATA 4957
Db |||||
QY 1902 TGTGTCTGAACTGGCTGTCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCC 1961
Db |||||
QY 4956 TGTGTCTCAACTGGCTTCTCAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCC 4897
Db |||||
QY 1962 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGAT 2021
Db |||||
QY 4896 TGTCTTTTAAACTGGCATCATTTCTGTGTAAAGCACACTTGAAGACAAGTACAGAT 4837
Db |||||
QY 2022 ACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCGCAGCGCTGGCCTCC 2081
Db |||||
QY 4836 ACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCGCAGCGCTGGCCTCC 4777
Db |||||
QY 2082 TTCTGTCATGATTCTATCCAA 2101
Db |||||
QY 4776 TTCTGTCATGATTCTATTCAA 4757
Db |||||

RESULT 5
US-09-091-501B-7
: Sequence 7, Application US/09091501B
: Patent No. 6518413
: GENERAL INFORMATION:
: APPLICANT: Tinsley, Jonathon M
: APPLICANT: Davies, Kay E
: TITLE OF INVENTION: Utrrophin gene expression
: FILE REFERENCE: 620-42
: CURRENT APPLICATION NUMBER: US/09/091,501B
: CURRENT FILING DATE: 1998-06-18
: PRIOR APPLICATION NUMBER: PCT/GB96/03156
: PRIOR FILING DATE: 1996-12-19
: PRIOR APPLICATION NUMBER: GB 9525962.8
: PRIOR FILING DATE: 1995-12-19
: PRIOR APPLICATION NUMBER: GB 9615797.9
: PRIOR FILING DATE: 1996-07-26
: PRIOR APPLICATION NUMBER: GB 9622174.2
: PRIOR FILING DATE: 1996-10-24
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 7
: LENGTH: 6045

: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (11)..(6037)
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Chimeric
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: (724)..(758)
: OTHER INFORMATION: Precise residue is left open
US-09-091-501B-7

Query Match 26.1%; Score 547.4; DB 4; Length 6045;
Best Local Similarity 61.4%; Pred. No. 5.2e-159;
Matches 900; Conservative 0; Mismatches 556; Indels 9; Gaps 1;
QY 645 ACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAAGAAC 704
Db |||||
QY 3270 ACAGCAGAAAGATGGTAAAGCTTTGGGAAATTTCTGAAGAGGCTACTATGCTTCAACATC 3329
Db |||||
QY 705 AACTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGCCGCTGGG 764
Db |||||
QY 3330 GACTGGATGATATGAACCAAGATGGAATGACTTAAAGCAAAATCTGCTAGCATCAGGG 3389
Db |||||
QY 765 TTTCTTTTACAAGACAGTTTCTGACCACTGGAAGCTGTGCACTTTCTCTGCAAGAACTTC 824
Db |||||
QY 3390 CCCATTTGGAGGCCAGCTGAGAAAGTGAACAGGTTGCTGATGTCTTTAGAAAGAACTGA 3449
Db |||||
QY 825 TGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACT 884
Db |||||
QY 3450 TCAATGGCTGAATATGAAGATGAAGAGCTTAAGAAACAAATGCCTATTGGAGGAGATG 3509
Db |||||
QY 885 TTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGCGCCTTCAAGAGGGAATTAAGAACTA 944
Db |||||
QY 3510 TTCCAGCCTTACAGCTCCAGTATGACCATTTGTAAGGCCCTGAGACGGGAGTTAAAGGAGA 3569
Db |||||
QY 945 AAGAACTTGTAAATCATGAGTACTCTTTGAGACTGTGACGAATATTTCTGACAGAGCAGCCTT 1004
Db |||||
QY 3570 AAGAAATATTCTGTCTCTGAATGCTGTGACCCAGSCCGAGTTTTCTTGGCTGATCAGCCAA 3629
Db |||||
QY 1005 T-----GGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCTCTCTGAGG 1055
Db |||||
QY 3630 TTGAGGCCCTTGAAGAGCCCAAGAAACCTTACAAATCAAAAACAGAAATTAACCTCTGAGG 3689
Db |||||
QY 1056 AGAGAGCCCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGT 1115
Db |||||
QY 3690 AGAGAGCCCCAAAAGATGCCAAAGCCATGCCAAACAGTCTTCTGAAGTCAAAAGAAAAT 3749
Db |||||
QY 1116 GGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAAATAGATGAGACCCTTGAAA 1175
Db |||||
QY 3750 GGGAAAGTCTAAATGCTGTAACTAGCAATTGGCAAAAGCAAGTGGAACAGCATTTGGAGA 3809
Db |||||
QY 1176 GACTCCAGGAACCTTCAAGAGGCCACCGATGAGTGGACCTCAAGCTGGCCCAAGCTGAGG 1235
Db |||||
QY 3810 AACTCAGAGACCTGACGGAGCTATGGATGACCTGGACGCTGACATGAAGAGGCAGAGT 3869
Db |||||
QY 1236 TGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCTCATTTGAGTCTCTCAAGATCACC 1295
Db |||||
QY 3870 CGGTGCGGAATGGCTGGAAGCCCGTGGGAGACTTACTCATTGACTCGCTGAGGATCACA 3929
Db |||||
QY 1296 TCGAGAAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCTCTGAAAAGAGAACTGAGCCACG 1355
Db |||||
QY 3930 TTGAAAAAATCATGGCATTTAGAGAAGAAATTGCAACCAATCAACTTTAAAGTTAAAAACGG 3989
Db |||||
QY 1356 TCAATGACCTTGCTCGCAGCTTACCACCTTTTGGGCACTTCAGCTCTCACCGTATAACCTCA 1415
Db |||||
QY 3990 TGAATGATTTATCCAGTCAGCTGTCTCCACITGACCTGCATCCCTCTCTTAAAGATGTCTC 4049
Db |||||
QY 1416 GCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAG 1475
Db |||||
QY 4050 GCCAGCTAGATGACCTTAATATGCGATGGAAACTTTTACAGGTTTCTGTGATGATCGCC 4109
Db |||||

Qy 1476 TCAGGAGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTCTTTCCA 1535
Db 4110 TTAACAGCTTCAGGAAGCCACAGAGATTTTGGACCATCTCTCAGCACTTCTCTCTA 4169
Qy 1536 CGTCTGTCCAGGGTCCCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCA 1595
Db 4170 CGTCAGTCCAGCTGCCGTGGCAAAAGATCCATTTACATAATAAAGTGCCCTATTACATCA 4229
Qy 1596 ACCACGAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTTACCAGTCTT 1655
Db 4230 ACCATCAAAACACAGACCACCTGTTGGGACCATCTTAAATGACCGAATCTTCAATCCC 4289
Qy 1656 TAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGAC 1715
Db 4290 TTGCTGACCTGAATAATGACGTTTCTGCTACCGGTACAGCAATCAAAATCCGAAGAC 4349
Qy 1716 TGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCTTGGACC 1775
Db 4350 TACAAAAAGCACTATGTTTGGATCTCTTAGAGTTGAGTACAAACAAATGAAATTTTCAAAC 4409
Qy 1776 AGCACAACCTCAAGCAAAATGACCAGCCCATGGATATCTCTGCAGATTATTAATTGTTGA 1835
Db 4410 AGCACAAGTTGAACCAAAATGACCAGCTCCTCAGTGTCCAGATGTCACTCAACTGTCTGA 4469
Qy 1836 CCACTATTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCG 1895
Db 4470 CAACAACCTTATGATGGACTTGAGCAAAATGCATAAGGACCTGGTCAACGTTCCACTCTGTG 4529
Qy 1896 TGGATATGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCC 1955
Db 4530 TTGATATGTGTCTCAATGGTTGCTCAATGTCTATGACACGGGTGGAACGTTGAAATTA 4589
Qy 1956 GTGCTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGT 2015
Db 4590 GAGTGCAGAGTCTGAAGATGGATTAAATGTCTCTCTCCAAAGTCTCTTGGAAAGAAAAT 4649
Qy 2016 ACAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGGAGGCTGG 2075
Db 4650 ACAGATATCTCTTTAAGGAAGTTGCGGGCGGACAGAAAATGTGTGACCAAGGAGGAGCTGG 4709
Qy 2076 GCCTCCTTCTGCATGATTTCTATCCA 2100
Db 4710 GCCTGTACTTTCATGATGCCATCCA 4734

RESULT 6
US-09-091-501B-9
; Sequence 9, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 10320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(10312)

FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Full length
; OTHER INFORMATION: utrophin construct
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (724)..(758)
; OTHER INFORMATION: Precise residue is left open
US-09-091-501B-9
Query Match 26.1%; Score 547.4; DB 4; Length 10320;
Best Local Similarity 61.4%; Pred. No. 7.5e-159;
Matches 900; Conservative 0; Mismatches 556; Indels 9; Gaps 1;
Qy 645 ACATGGTGGTGTAGTTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGGAGAAGAC 704
Db 7545 ACAGGCAGAAGATGTTAAAAGCTTTTGGGAAATTTCTGAAGAGGCTACTATGCTTCAACATC 7604
Qy 705 AACTTAAGGTATTGGAGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGCTGGG 764
Db 7605 GACTGGATGATATGAACCAAGATGGAATGACTTAAAGCAAAATCTGCTAGCATCAGGG 7664
Qy 765 TTCTTTTACAAGACAGTCTTGACCAGTGGAAAGCTCTGCACCTTTTCTCTGCAGGAACCTC 824
Db 7665 CCCATTGGAGGCCAGCGCTGAGAAGTGGAAACAGGTTGCTGATGTCCTTAGAAGAACTGA 7724
Qy 825 TGGTGTGGCTACAGCTGAAAAGATGATGAATTAAGCCGAGGACCTATTGGAGGCGACT 884
Db 7725 TCAAATGGCTGAATATGAAAAGATGAAGAGCTTAAGAAAACAAATGCCTATTGGAGGAGATG 7784
Qy 885 TTCCAGCAGTTTCAGAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTTGAAAACATA 944
Db 7785 TTCCAGCCTTACAGCTCCAGTATGACCAATTGTAAGGCCCTGAGACGGGAGTTAAAGGAGA 7844
Qy 945 AAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTT 1004
Db 7845 AAGAATATTCTGCTCTGAATGCTGTCGACCAGGCCCGAGTTTCTTGGCTGATCAGCCAA 7904
Qy 1005 T-----GGAGGACTAGAGAACTCTACAGAGAGCCCGAGAGAGCTGCCTCCTGAGG 1055
Db 7905 TTGAGGCCCTTGAAGAGCCAAAGAAACCTTACAAATCAAAAACAGAAATTAACCTCCTGAGG 7964
Qy 1056 AGAGAGCCCGAGAAATGTCACCTGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGT 1115
Db 7965 AGAGAGCCCAAAAGATTGCCAAAGCCATGCGCAAAACAGTCTTCTGAAGTCAAAAGAAAAT 8024
Qy 1116 GGGAAAAATTGAACCTGCACCTCCGCTGACTGSCAGAGAAAAATAGATGAGACCCCTTGAAA 1175
Db 8025 GGGAAAGTCTAAATGCTGTAACTAGCAATTTGGCAAAAGCAAGTGGACAGGCAATTTGGAGA 8084
Qy 1176 GACTCCAGGAACCTTCAAGAGGCCACCGATGAGTGGACCTCAAGCTGCGCCAAGCTGAGG 1235
Db 8085 AACTCAGAGACCTGACAGGAGCTATGGATGACCTGGACGCTGACATGAGAGGAGGAGT 8144
Qy 1236 TGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCATTGACTCTCTCCAAGATCACC 1295
Db 8145 CCCTGCGGAATGGCTGGAAAGCCCGTGGGAGACTTACTCATTGACTCGCTGAGGATCACA 8204
Qy 1296 TCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCCCTCTGAAAAGAGACGCTGAGCCACG 1355
Db 8205 TTGAAAAAATCATGGCATTTAGAGAAGAAATTTGCACCAATCAA CTTTAAAGTTAAACGG 8264
Qy 1356 TCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCA 1415
Db 8265 TGAATGATTTATCCAGTCAGCTGTCTCCACTTGAACCTGTCATCCCTCTCTTAAAGATGTCTC 8324
Qy 1416 GCACTCTGGAAGACCTGAACACAGATGGAAGCTTCTGAGGTGGCCGTGAGGACCGAG 1475
Db 8325 GCCAGCTAGATGACCTTAATATATGCGATGGAAGAACTTTTACAGGTTTCTGTGGATGATCGCC 8384
Qy 1476 TCAGGCAGCTGCATGAAGCCACAGGAGCTTTTGGTCCAGCATCTCAGCACTTCTTTCCA 1535
Db 8385 TTAACACAGCTTCAGGAAGCCACAGAGATTTTGGACCATCCTCTCAGCACTTCTCTCTA 8444

| | | | |
|----|------|---|------|
| 2Y | 1536 | CGTCTGTCCAGGTCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCA | 1595 |
| | | | |
| 2b | 8445 | CGTCAGTCCAGTGCCTGGCAAAGATCCATTTCACATATATAAGTGCCCTATTACATCA | 8504 |
| | | | |
| 2Y | 1596 | ACCACGAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTTT | 1655 |
| | | | |
| 2b | 8505 | ACCATCAAAACACAGACCACCTGTTGGGACCATCTATAAATGACCGAACTCTTTCAATCCC | 8564 |
| | | | |
| 2Y | 1656 | TAGCTGACCTGGAATAATGTGAGATTCTCAGCTTATAGGACTGTCATGAACTCCGAAGAC | 1715 |
| | | | |
| 2b | 8565 | TTGCTGACCTGAATAATGTACGTTTCTGCCTACCGTACAGCAATCAAAATCCGAAGAC | 8624 |
| | | | |
| 2Y | 1716 | TGCAGAAGGCCCTTTGCTTTGGATCTCTTGAGCCTGTGAGCTGATGTGATGCCCTTGACC | 1775 |
| | | | |
| 2b | 8625 | TACAAAAGCACTATGTTTGGATCTCTTAGAGTTTGAGTACAACAAAATGAAATTTTCAAAC | 8684 |
| | | | |
| 2Y | 1776 | AGCACAACTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAAATGTTTGA | 1835 |
| | | | |
| 2b | 8685 | AGCACAAAGTTGAACCAAAATGACCAGCTCCTCAGTGTTCCAGATGTCTATCAACTGTCTGA | 8744 |
| | | | |
| 2Y | 1836 | CCACTATTTATGACCGCCTGGAGCAAGAGACACAACAATTTGGTCAACGTCCTCTCTGCG | 1895 |
| | | | |
| 2b | 8745 | CAACAACTTATGATGGACTTGAGCAAAATGCATAAGGACCTGGTCAACGTTCCACTGTGTG | 8804 |
| | | | |
| 2Y | 1896 | TGGATATGTCTGAACCTGGCTGCTGAAATGTTTATGATACGGGACGAACAGGGAGGATCC | 1955 |
| | | | |
| 2b | 8805 | TTGATATGTGTCTCAATTGGTTGCTCAATGTCTATGACACGGGTGGAACCTGGAAAAATTA | 8864 |
| | | | |
| 2Y | 1956 | GTGTCCTGCTTTTAAAACTGGCATCATTTCCCTGTGTATAAGCACATTTGGAAGACAAAGT | 2015 |
| | | | |
| 2b | 8865 | GAGTGCAGAGTCTGAAGATTGGATTAAATGTCTCTCTCCAAAAGGTCTCTTTGGAAGAAAAAT | 8924 |
| | | | |
| 2Y | 2016 | ACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGCAGGCTGG | 2075 |
| | | | |
| 2b | 8925 | ACAGATATCTCTTTAAGGAAGTTGCGGGGCCGACAGAAAATGTGTGACCAAGAGGCGACTGG | 8984 |
| | | | |
| 2Y | 2076 | GCCTCCTTCTGCATGATTCATCCA | 2100 |
| | | | |
| 2b | 8985 | GCCTGTTACTTTCATGATGCCATCCA | 9009 |
| | | | |

RESULT 7

IS-09-091-501B-5
Sequence 5, Application US/09091501B
Patent No. 6518413
GENERAL INFORMATION:
APPLICANT: Tinsley, Jonathon M
APPLICANT: Davies, Kay E
TITLE OF INVENTION: Utrrophin gene expression
FILE REFERENCE: 620-42
CURRENT APPLICATION NUMBER: US/09/091,501B
CURRENT FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: PCT/GB96/03156
PRIOR FILING DATE: 1996-12-19
PRIOR APPLICATION NUMBER: GB 9525962.8
PRIOR FILING DATE: 1995-12-19
PRIOR APPLICATION NUMBER: GB 9615797.9
PRIOR FILING DATE: 1996-07-26
PRIOR APPLICATION NUMBER: GB 9622174.2
PRIOR FILING DATE: 1996-10-24
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1

Query Match 3.8%; Score 79.4; DB 4; Length 200;
Best Local Similarity 64.3%; Pred. No. le-14;
Matches 119; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
y 565 CCTAAACGCCAAGTACAACAACATAAGGTGCTTCAAGAGAGATCTAGAACAAAGACAAGT 624

| | | | | | |
|----|-----|---|---------|--------------------|-----|
| Db | 16 | CCTGCAAAACCTGCTTGAAGAACATAAAGTTTGC | AAAGTGA | CCTCGAAGCTGAGCAGGT | 75 |
| QY | 625 | CAGGGTCAATTCCTCACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGC | | | 684 |
| Db | 76 | GAAGGTGAATTCCTTAACTCATATGGTGGTGATTGTGGATGAAAAACAGTGGGGAGAGCGC | | | 135 |
| QY | 685 | AACTGCTGCTTTGGAAAGAACAACTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAG | | | 744 |
| Db | 136 | CACAGCTGTTTTGGAAAGATCAGTTACAGAAACTGGGTGAGCGCTGGACAGCTGTATGCCG | | | 195 |
| QY | 745 | ATGGA | 749 | | |
| Db | 196 | CTGGA | 200 | | |

RESULT 8

```

US-09-091-501B-4
; Sequence 4, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathan M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 4
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Mus sp.
US-09-091-501B-4

```

Query Match 3.7%; Score 78.6; DB 4; Length 200;
Best Local Similarity 62.4%; Pred. No. 1.8e-14;
Matches 123; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

| | | | |
|----|-----|---|-----|
| QY | 553 | TGATCTTTGAAGACCTTAAACCGCCAAAGTACAAACAATAAAGGTGCTTCAAGAAGATCTAGA | 612 |
| | | | |
| Db | 4 | TGACCTGCCCTCCCTGCAGAAGCTGCTTCAAGAACATATAAAGTTTGCAAAATGACCTTGA | 63 |
| | | | |
| QY | 613 | ACAAGAACBAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGTTTGATGAATCTAG | 672 |
| | | | |
| Db | 64 | AGCTGAACACAGGTGAAGGTAAATTCCTTAACTCACATGGTGGTGATTGTGGATGAACAACAG | 123 |
| | | | |
| QY | 673 | TGGAGATCACGCAACTGCTGCTTTTGGAGAGAACAACTTAAGGTATTGGGAGATCGATGGGC | 732 |
| | | | |
| Db | 124 | TGGGGAGAGTGCACAGCTCTTCTGGAAGATCAGTTACAGAACTGGGTGAGCGCTGGAC | 183 |
| | | | |
| QY | 733 | AAACATCTGTAGATGGA | 749 |
| | | | |
| Db | 184 | AGCTGTATGCCGCTGGA | 200 |
| | | | |

RESULT 9

RESULT 9
 US-09-091-501B-6
 ; Sequence 6, Application US/09091501B
 ; Patent No. 6518413
 ; GENERAL INFORMATION:
 ; APPLICANT: Tinsley, Jonathon M
 ; APPLICANT: Davies, Kay E
 ; TITLE OF INVENTION: Urothrin gene expression
 ; FILE REFERENCE: 620-42

NAME/KEY: misc feature
LOCATION: (570001)..(585000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (585001)..(600000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (600001)..(615000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (615001)..(630000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (630001)..(645000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (645001)..(660000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (660001)..(675000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (675001)..(690000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (690001)..(705000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (705001)..(720000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (720001)..(735000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (735001)..(750000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (750001)..(765000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (765001)..(780000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (780001)..(795000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (795001)..(810000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (810001)..(825000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (825001)..(840000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (840001)..(855000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (855001)..(870000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (870001)..(885000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (885001)..(900000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (900001)..(915000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature

Query Match 2.1%; Score 44; DB 4; Length 1230025;
Best Local Similarity 53.5%; Pred. No. 0.35;

Matches 92; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
QY 342 AATTATCAGAGATGAAGAAAACCTGAAGTACAGAGCAGATGAATCTCTAAATTCAGAT 401
|||||
Db 656477 AATTAGAAGAGAGAGAGAAAGAGAAATTGAGGATATCAAAGACTCAGATACAAAT 656418
|||||
QY 402 GGAATGCCTCAGGCTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGG 461
|||||
Db 656417 GGGTTTCGATCACTCAAGCTGCTAAATTACATAACGTCACTAGGCAAGCAATTTATGTGG 656358
|||||
QY 462 ATCTCCAGAATCAGAAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGA 513
|||||
Db 656357 CAATTAAGCAGAAAAAACTAAAAGCTTCTAAAGAGACGCGCTGGGAAATAGA 656306
|||||
RESULT 13
US-09-107-532A-1186
; Sequence 1186, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 1186:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1179 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...1179
; SEQUENCE DESCRIPTION: SEQ ID NO: 1186:
US-09-107-532A-1186

Query Match 2.1%; Score 43.4; DB 4; Length 1179;
Best Local Similarity 50.7%; Pred. No. 0.005;
Matches 104; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 471 ATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAACAGGAAAA 530

Db 602 AGCAAGAATAAAGATTGATCGATCAGACAAAGAAAAATGGAGATACGATCGGAGAA 661
2Y 531 TGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAAGCGCAAGTACAAACATA 590
Db 662 TTGTAGAAGTGCTCGTTGGAGCGGTTCCAGCTGGATTAGGAAGCTACGTACAAATGGACA 721
2Y 591 AGGTGCTTCAAGAAGATCTAGAACAAAGAACAGTCAAGGTCATCTCTCACTCACATGG 650
Db 722 CGAAGCTAGATGCCAAATCGCACAAAGCTGTGGTTAGTATCAATGCCCTTTAAAGCGTAG 781
2Y 651 TGCTGTAGTTGATGAATCTAGTGG 675
Db 782 AATTGGGGTCGGATTCACTTCTGG 806

RESULT 14
US-09-620-312D-69
; Sequence 69, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 69
LENGTH: 1690
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (128)..(1522)
US-09-620-312D-69

Query Match 2.0%; Score 42.8; DB 4; Length 1690;
Best Local Similarity 47.9%; Pred. No. 0.0098;
Matches 156; Conservative 0; Mismatches 167; Indels 3; Gaps 1;
Y 1180 CCAGGAACCTTCAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT 1239
b 55 CAAGGAGTTGCACCAAGTGGCGCACGACTGGACGACGAGCTGGCATGGGTTTCAGGAGCG 114
Y 1240 CAAGGGA---TCCTGGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCT 1296
b 115 GCTGCCACTGGCCATGCAGACAGAGCGGCAACGGTTTGCAGGCGGTCCAGCAGACAT 174
Y 1297 CGAGAAAGTCAAGCACTTCGAGGAGAGAAATTGCGCCTCTGAAAGAGAACGTTGAGCCCGT 1356
b 175 CAAAAGAAACCAAGGCGCTGCGGGGGGAGATCCAGGGCGCATGGGCGCGCTTGAGGAGGT 234

QY 1357 CAATGACCTTGCTGCGCCAGCTTACCACATTTGGGCATTACGCTCTCACCGTATAACCTCAG 1416
Db 235 GCTGGAGCGCGCGGCGGTGGCTGGCTGCGCAGCCGAGGAGGAGGAGGAGGAGGAGGAGG 294
QY 1417 CACTCTGGAAGACCTGAACACACAGATGGAAGCTTTTCAGAGTGGCCGTCGAGGACCGAGT 1476
Db 295 GGGCCTGGAGCAGCTGCAGAGCGCCTGGCCGCGGACTGCGGAGGCTGCCGAGCGACGCA 354
QY 1477 CAGGACGCTGCATGAAGCCACAGGG 1502
Db 355 GCAGGTGCTGGACGCCGCCCTTCCAGG 380

RESULT 15
US-09-368-590-1
; Sequence 1, Application US/09368590
; Patent No. 6187563
; GENERAL INFORMATION:
; APPLICANT: Solimena, Michele
; TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
; TITLE OF INVENTION: AUTOANTIGENS OF AUTOIMMUNE DISEASES
; FILE REFERENCE: 101918-200 (OCR-941)
; CURRENT APPLICATION NUMBER: US/09/368,590
; CURRENT FILING DATE: 1999-08-04
; EARLIER APPLICATION NUMBER: 60/095,657
; EARLIER FILING DATE: 1998-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 7812
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(6879)
; NAME/KEY: unsure
; LOCATION: (100)...(102)
; NAME/KEY: unsure
; LOCATION: (1021)...(1023)
; NAME/KEY: unsure
; LOCATION: (2266)...(2268)
US-09-368-590-1

Query Match 2.0%; Score 42.8; DB 3; Length 7812;
Best Local Similarity 47.9%; Pred. No. 0.027;
Matches 156; Conservative 0; Mismatches 167; Indels 3; Gaps 1;
QY 1180 CCAGGAACCTTCAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 1239
Db 3639 CAAGGAGTTGCACAGGTGGCGCACGACCTGGACGAGCTGSCATGGGTTTCAGGAGCG 3698
QY 1240 CAAGGGA---TCCTGGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCT 1296
Db 3699 GCTGCCACTGGCCATGCAGACAGAGCGGCAACGGTTTGCAGGCGGTCCAGCAGACAT 3758
QY 1297 CGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTTGAGCCACGT 1356
Db 3759 CAAAAGAACCCAGGCGCTGCGGCGGGAGATCCAGGCGCATGGGCGCGCTTGAGGAGGT 3818
QY 1357 CAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAACCTCAG 1416
Db 3819 GCTGGAGCGCGGGCGCTGGCGTCTGCTGCGCAGCCCGGAGGAGGAGGAGGAGGAGGAG 3878
QY 1417 CACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGAGTGGCCGCTCGAGGACCGAGT 1476
Db 3879 GGGCCTGGAGCAGCTGCAGAGCGCTGGGCGGACTGCGGAGGAGGAGGAGGAGGAGGAG 3938
QY 1477 CAGGACGCTGCATGAAGCCACAGGG 1502
Db 3939 GCAGGTGCTGGACCGCGCTTCCAGG 3964

Search completed: February 2, 2004, 06:16:01
Job time : 133.879 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 23:51:42 ; Search time 687.863 Seconds
(without alignments)
11131.886 Million cell updates/sec

Title: US-09-845-416-12_COPY_900_3000
Perfect score: 2101
Sequence: 1 gagctatgcctacacacagg.....ttctgcatgattctatccaa 2101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2434939 seqs, 1822278265 residues

Total number of hits satisfying chosen parameters: 4869878

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|--------|---------------|--------|-------|-------------------|
| 1 | 2101 | 100.0 | 3510 | 13 | US-09-845-416-12 |
| 2 | 2101 | 100.0 | 4476 | 13 | US-09-845-416-31 |
| 3 | 1760.4 | 83.8 | 5339 | 13 | US-10-149-736-40 |
| 4 | 1750.8 | 83.3 | 3531 | 13 | US-09-845-416-10 |
| 5 | 1750.8 | 83.3 | 4498 | 13 | US-09-845-416-30 |
| 6 | 1743 | 83.0 | 3858 | 13 | US-09-845-416-9 |
| 7 | 1743 | 83.0 | 4825 | 13 | US-09-845-416-29 |
| 8 | 1743 | 83.0 | 4848 | 13 | US-09-845-416-35 |
| 9 | 1743 | 83.0 | 5060 | 13 | US-09-845-416-36 |
| 10 | 1641 | 78.1 | 4414 | 13 | US-09-845-416-32 |
| 11 | 1629 | 77.5 | 3446 | 13 | US-09-845-416-14 |
| 12 | 1629 | 77.5 | 5462 | 13 | US-10-149-736-41 |
| 13 | 1611.6 | 76.7 | 5417 | 13 | US-10-149-736-39 |
| 14 | 1602 | 75.2 | 3999 | 13 | US-09-845-416-6 |
| 15 | 1602 | 75.2 | 4966 | 13 | US-09-845-416-28 |
| | | | | | Sequence 12, Appl |
| | | | | | Sequence 31, Appl |
| | | | | | Sequence 40, Appl |
| | | | | | Sequence 10, Appl |
| | | | | | Sequence 30, Appl |
| | | | | | Sequence 9, Appl |
| | | | | | Sequence 29, Appl |
| | | | | | Sequence 35, Appl |
| | | | | | Sequence 36, Appl |
| | | | | | Sequence 32, Appl |
| | | | | | Sequence 14, Appl |
| | | | | | Sequence 41, Appl |
| | | | | | Sequence 39, Appl |
| | | | | | Sequence 6, Appl |
| | | | | | Sequence 28, Appl |

| | | | | | | |
|----|--------|------|-------|----|---------------------|-------------------|
| 16 | 1602 | 76.2 | 4990 | 13 | US-09-845-416-34 | Sequence 34, Appl |
| 17 | 1419 | 67.5 | 4182 | 13 | US-09-845-416-2 | Sequence 2, Appl |
| 18 | 1419 | 67.5 | 5149 | 13 | US-09-845-416-27 | Sequence 27, Appl |
| 19 | 1328.8 | 63.2 | 2169 | 13 | US-09-845-416-4 | Sequence 4, Appl |
| 20 | 1328.8 | 63.2 | 8689 | 13 | US-10-149-736-42 | Sequence 42, Appl |
| 21 | 1328.8 | 63.2 | 11058 | 13 | US-09-845-416-1 | Sequence 1, Appl |
| 22 | 1328.8 | 63.2 | 11443 | 13 | US-10-149-736-44 | Sequence 44, Appl |
| 23 | 1328.8 | 63.2 | 12057 | 13 | US-10-149-736-47 | Sequence 47, Appl |
| 24 | 1328.8 | 63.2 | 13957 | 10 | US-09-782-378A-22 | Sequence 22, Appl |
| 25 | 1328.8 | 63.2 | 13957 | 10 | US-09-880-107-2284 | Sequence 2284, Ap |
| 26 | 1328.8 | 63.2 | 13957 | 13 | US-10-149-736-1 | Sequence 1, Appl |
| 27 | 1328.8 | 63.2 | 14082 | 13 | US-10-341-434-108 | Sequence 108, App |
| 28 | 1324 | 63.0 | 1821 | 13 | US-09-845-416-13 | Sequence 13, Appl |
| 29 | 1149.6 | 54.7 | 13815 | 13 | US-10-149-736-2 | Sequence 2, Appl |
| 30 | 937 | 44.6 | 1434 | 13 | US-09-845-416-15 | Sequence 15, Appl |
| 31 | 785.4 | 37.4 | 1991 | 13 | US-09-845-416-3 | Sequence 3, Appl |
| 32 | 777 | 37.0 | 1667 | 13 | US-09-845-416-7 | Sequence 7, Appl |
| 33 | 554.2 | 26.4 | 11096 | 13 | US-10-149-736-4 | Sequence 4, Appl |
| 34 | 547.4 | 26.1 | 10302 | 10 | US-09-782-378A-23 | Sequence 23, Appl |
| 35 | 547.4 | 26.1 | 10302 | 13 | US-10-149-736-3 | Sequence 3, Appl |
| 36 | 538.6 | 25.6 | 16531 | 13 | US-10-101-510-667 | Sequence 667, App |
| 37 | 450 | 21.4 | 1340 | 13 | US-09-845-416-11 | Sequence 11, Appl |
| 38 | 393 | 18.7 | 887 | 13 | US-10-149-736-35 | Sequence 35, Appl |
| 39 | 387 | 18.4 | 387 | 13 | US-10-149-736-32 | Sequence 32, Appl |
| 40 | 331 | 15.8 | 333 | 13 | US-10-149-736-9 | Sequence 9, Appl |
| 41 | 327 | 15.6 | 327 | 13 | US-10-149-736-8 | Sequence 8, Appl |
| 42 | 324 | 15.4 | 324 | 13 | US-10-149-736-33 | Sequence 33, Appl |
| 43 | 216 | 10.3 | 216 | 13 | US-10-149-736-34 | Sequence 34, Appl |
| 44 | 190.2 | 9.1 | 256 | 9 | US-09-864-761-21956 | Sequence 21956, A |
| 45 | 178 | 8.5 | 466 | 9 | US-09-864-761-6092 | Sequence 6092, Ap |

ALIGNMENTS

RESULT 1

US-09-845-416-12
; Sequence 12, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 3510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-12

| | | | | | |
|-----------------------|------|--|---------------|-----------|--------------|
| Query Match | | 100.0%; | Score 2101; | DB 13; | Length 3510; |
| Best Local Similarity | | 100.0%; | Pred. No. 0; | | |
| Matches 2101; | | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| QY | 1 | GAGCTATGCCTACACACAGCTGCTTATGTGTCACCACTCTGACCCCTACACGGAGCCCAT | 60 | | |
| DB | 900 | GAGCTATGCCTACACACAGCTGCTTATGTGTCACCACTCTGACCCCTACACGGAGCCCAT | 959 | | |
| QY | 61 | TCCITCACAGCATTGGAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG | 120 | | |
| DB | 960 | TCCITCACAGCATTGGAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG | 1019 | | |
| QY | 121 | TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC | 180 | | |
| DB | 1020 | TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC | 1079 | | |
| QY | 181 | TGCTGAGGACACATTGCAAGACAAGGAGAGATTCTTCTAATGATGTGGAAGTGGTGAAGA | 240 | | |

SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 31
LENGTH: 4476
TYPE: DNA
ORGANISM: Homo sapiens
IS-09-845-416-31

Query Match 100.0%; Score 2101; DB 13; Length 4476;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|---|------|--|------|
| y | 1 | GAGCTATGCCTACACACAGGCTGCTTATGTCAACCCTCTGTGACCCCTACACGAGCCCAT | 60 |
| b | 1656 | GAGCTATGCCTACACACAGGCTGCTTATGTCAACCCTCTGTGACCCCTACACGAGCCCAT | 1715 |
| y | 61 | TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTCAATTGATGGAGAG | 120 |
| b | 1716 | TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTCAATTGATGGAGAG | 1775 |
| y | 121 | TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC | 180 |
| b | 1776 | TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAAGTATTATCGTGGCTTCTTTC | 1835 |
| y | 181 | TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAGA | 240 |
| b | 1836 | TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAGA | 1895 |
| y | 241 | CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG | 300 |
| b | 1896 | CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG | 1955 |
| y | 301 | TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATATATCAGAAGATGAAGA | 360 |
| b | 1956 | TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATATATCAGAAGATGAAGA | 2015 |
| y | 361 | AACTGAAGTACAAGACAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC | 420 |
| b | 2016 | AACTGAAGTACAAGACAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC | 2075 |
| y | 421 | TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT | 480 |
| b | 2076 | TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT | 2135 |
| y | 481 | GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAAGGAAAGTGAAGA | 540 |
| b | 2136 | GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAAGGAAAGTGAAGA | 2195 |
| y | 541 | GCCTCTTGGAACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAACATAAGGTGCTTCA | 600 |
| b | 2196 | GCCTCTTGGAACCTGATCTTGAAGACCTTAAACGCCCAAGTACAACAACATAAGGTGCTTCA | 2255 |
| y | 601 | AGAAGATCTAGAACAAAGCAAGTCAAGGTCGAATTTCTCACTACATGATGGTGGTAGT | 660 |
| b | 2256 | AGAAGATCTAGAACAAAGCAAGTCAAGGTCGAATTTCTCACTACATGATGGTGGTAGT | 2315 |
| y | 661 | TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAAACAACTTAAGGTATTGGG | 720 |
| b | 2316 | TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAAACAACTTAAGGTATTGGG | 2375 |
| y | 721 | AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGCCGCTGGGTTCTTTTACAAGACAG | 780 |
| b | 2376 | AGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGCCGCTGGGTTCTTTTACAAGACAG | 2435 |
| y | 781 | TTCTGACCAGTGGAGCGTCTGCACCTTTCTCTGCAGGAACCTTCGGTGTGGCTACAGCT | 840 |
| b | 2436 | TTCTGACCAGTGGAGCGTCTGCACCTTTCTCTGCAGGAACCTTCGGTGTGGCTACAGCT | 2495 |
| y | 841 | GAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAA | 900 |
| b | 2496 | GAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAA | 2555 |
| y | 901 | GCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACTAAGAAACCTGTAATCAT | 960 |

| | | | |
|----|------|--|------|
| Db | 2556 | GCAGAAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCAT | 2615 |
| Qy | 961 | GAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAA | 1020 |
| Db | 2616 | GAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAA | 2675 |
| Qy | 1021 | ACTCTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCGAGAAATGTCACTCGGCT | 1080 |
| Db | 2676 | ACTCTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCGAGAAATGTCACTCGGCT | 2735 |
| Qy | 1081 | TCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCACTCCGC | 1140 |
| Db | 2736 | TCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTGCACTCCGC | 2795 |
| Qy | 1141 | TGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCAC | 1200 |
| Db | 2796 | TGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCAC | 2855 |
| Qy | 1201 | GGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT | 1260 |
| Db | 2856 | GGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT | 2915 |
| Qy | 1261 | GGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGG | 1320 |
| Db | 2916 | GGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGG | 2975 |
| Qy | 1321 | AGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTAC | 1380 |
| Db | 2976 | AGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTAC | 3035 |
| Qy | 1381 | CACCTTTGGGCAATCAGCTCTCACCGTATAACCTCAGCACTCTGGAGACCTGAACACCAG | 1440 |
| Db | 3036 | CACCTTTGGGCAATCAGCTCTCACCGTATAACCTCAGCACTCTGGAGACCTGAACACCAG | 3095 |
| Qy | 1441 | ATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGAGCTGTCCAGGTCCTGGGAGAG | 1500 |
| Db | 3096 | ATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGAGCTGTCCAGGTCCTGGGAGAG | 3155 |
| Qy | 1501 | GGACTTTGGTCCAGCATCTCAGCACCTTCTTTCCACGTCTGTCCAGGTCCTGGGAGAG | 1560 |
| Db | 3156 | GGACTTTGGTCCAGCATCTCAGCACCTTCTTTCCACGTCTGTCCAGGTCCTGGGAGAG | 3215 |
| Qy | 1561 | AGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCAGAGACTCAAAACAACTTGCTG | 1620 |
| Db | 3216 | AGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCAGAGACTCAAAACAACTTGCTG | 3275 |
| Qy | 1621 | GGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAAATAATGTGAGATT | 1680 |
| Db | 3276 | GGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAAATAATGTGAGATT | 3335 |
| Qy | 1681 | CTCAGCTTATAGGACTGCCATGAAACTCCGAAAGACTGCAGAAAGGCCCTTTGCTTGATCT | 1740 |
| Db | 3336 | CTCAGCTTATAGGACTGCCATGAAACTCCGAAAGACTGCAGAAAGGCCCTTTGCTTGATCT | 3395 |
| Qy | 1741 | CTTGAGCCTGTGAGCTGCATGTGATGCCTTGACCCAGCACAACTCAAGCAAAATGACCA | 1800 |
| Db | 3396 | CTTGAGCCTGTGAGCTGCATGTGATGCCTTGACCCAGCACAACTCAAGCAAAATGACCA | 3455 |
| Qy | 1801 | GCCCATGGATATCCTGCAGATTATTATTGTTTGACCACTATTTATGACCGCCTGGAGCA | 1860 |
| Db | 3456 | GCCCATGGATATCCTGCAGATTATTATTGTTTGACCACTATTTATGACCGCCTGGAGCA | 3515 |
| Qy | 1861 | AGAGCACAACTTTGGTCAACGTCCCTCTCTGCGTGGATATGTGTGAACTGGCTGCT | 1920 |
| Db | 3516 | AGAGCACAACTTTGGTCAACGTCCCTCTCTGCGTGGATATGTGTGAACTGGCTGCT | 3575 |
| Qy | 1921 | GAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCTGTCTTTTAAACTGGCAT | 1980 |
| Db | 3576 | GAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCTGTCTTTTAAACTGGCAT | 3635 |
| Qy | 1981 | CATTTCCCTGTGTAAAGCACATTTTGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGC | 2040 |
| Db | 3636 | CATTTCCCTGTGTAAAGCACATTTTGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGC | 3695 |

| | | | |
|----------------|------|--|------|
| Qy | 2041 | AAGTTCAACAGGATTTGTGACCAGCGCAGGCTGGGCCTCCTTCTGTCATGATTCTATCCA | 2100 |
| D _b | 3696 | AAGTTCAACAGGATTTGTGACCAGCGCAGGCTGGGCCTCCTTCTGTCATGATTCTATCCA | 3755 |
| Qy | 2101 | A | 2101 |
| D _b | 3756 | A | 3756 |

RESULT 3

US-10-149-736-40
; Sequence 40, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40
; LENGTH: 5339
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-40

| | Query Match | 83.8%; | Score 1760.4; | DB 13; | Length 5339; |
|----|-----------------------|--|---------------|-----------------|--------------------|
| | Best Local Similarity | 90.3%; | Pred. No. 0; | | |
| | Matches 1910; | Conservative | 0; | Mismatches 191; | Indels 15; Gaps 2; |
| QY | 1 | GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT | 60 | | |
| Db | 1099 | GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACGGAGCCCAT | 1158 | | |
| QY | 61 | TCCTTCACAGCATTTGGAAAGCTCCTGAAGACAAGTCATTTTGGCAGATTTCATTGATGGAGAG | 120 | | |
| Db | 1159 | TCCTTCACAGCATTTGGAAAGCTCCTGAAGACAAGTCATTTTGGCAGATTTCATTGATGGAGAG | 1218 | | |
| QY | 121 | TGAAGTAAACCTGGACCCTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC | 180 | | |
| Db | 1219 | TGAAGTAAACCTGGACCCTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC | 1278 | | |
| QY | 181 | TGCTGAGGACACATTGCAAGACACAAGGAGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA | 240 | | |
| Db | 1279 | TGCTGAGGACACATTGCAAGACACAAGGAGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA | 1338 | | |
| QY | 241 | CCAGTTTCATACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTGG | 300 | | |
| Db | 1339 | CCAGTTTCATACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTGG | 1398 | | |
| QY | 301 | TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTTATCAGAAGATGAAGA | 360 | | |
| Db | 1399 | TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTTATCAGAAGATGAAGA | 1458 | | |
| QY | 361 | AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCCTCAGGGTAGC | 420 | | |
| Db | 1459 | AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCCTCAGGGTAGC | 1518 | | |
| QY | 421 | TAGCATGAAAAACAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT | 480 | | |
| Db | 1519 | TAGCATGAAAAACAAGCAATTTACATCATAGATTACTGCAACAGTTCCCCCTGGACCT | 1578 | | |
| QY | 481 | GAAGAGTTGAATGACTGGCTAAACAAAAACAGAAGAAAGCAAGAAAAATGGAGGAAGA | 540 | | |

| | | | |
|------|----|---|------|
| 1579 | Db | GGAAAAGTTTCTTGGCTTGCTTACAGAAGCTGAAACAACACTGCCAATGTCTCTACAGGATGC | 1633 |
| 541 | QY | GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAAACAACATAAGGTGCTTCA | 600 |
| 1639 | Db | TACCCGTAAAGGAAAGGCTCTCTAGAAGACTCCAAGGGAGTAAAGAGAGCTGATGAAACAATG | 1698 |
| 601 | QY | AGAGATCTAGAACAAGAACAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGGTAGT | 660 |
| 1699 | Db | GCAAGACCTCCAAGGTGAAATTGAAAGCTCACACAGATGTTTTATCACAACCTGGATGAAAA | 1758 |
| 661 | QY | TGATGA-----ATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACCT | 708 |
| 1759 | Db | CAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCACTGCTGTACAAAGACG | 1818 |
| 709 | QY | TAAG-- --TATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGT | 765 |
| 1819 | Db | TTTTGGATAACATGAACITTCAGTGGAGTGAACITTCGGA AAAAGTCTCTCAACATTAGGTC | 1878 |
| 766 | QY | TCCTTTTCAAGACAGTTCTGACCACTGGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCT | 825 |
| 1879 | Db | CCATTTGGAAGCCAGTTCTGACCACTGGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCT | 1938 |
| 826 | QY | GGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTT | 885 |
| 1939 | Db | GGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTT | 1998 |
| 886 | QY | TCCAGCAGTTTCAAGACGAGATGTACATAGGGCCTTCAAGAGGGGAATTGAAAACTAA | 945 |
| 1999 | Db | TCCAGCAGTTTCAAGACGAGATGTACATAGGGCCTTCAAGAGGGGAATTGAAAACTAA | 2058 |
| 946 | QY | AGAACTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGCAGAGCAGCCTTT | 1005 |
| 2059 | Db | AGAACTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGCAGAGCAGCCTTT | 2118 |
| 1006 | QY | GGAAGGACTAGAGAAACTCTACAGGAGCCCGCTGAGGAGCTGCCTCTCTGAGGAGAGACCCA | 1065 |
| 2119 | Db | GGAAGGACTAGAGAAACTCTACAGGAGCCCGAGAGCTGCCTCTCTGAGGAGAGACCCA | 2178 |
| 1066 | QY | GAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATATCTGAGTGGGAAAAATT | 1125 |
| 2179 | Db | GAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATATCTGAGTGGGAAAAATT | 2238 |
| 1126 | QY | GAACTTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAGAAGACTCCAGGA | 1185 |
| 2239 | Db | GAACTTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAGAAGACTCCAGGA | 2298 |
| 1186 | QY | ACTTCAAGAGGCCACGGATGAGCTTGGACCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGG | 1245 |
| 2299 | Db | ACTTCAAGAGGCCACGGATGAGCTTGGACCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGG | 2358 |
| 1246 | QY | ATCCTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAAGT | 1305 |
| 2359 | Db | ATCCTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAAGT | 2418 |
| 1306 | QY | CAAGGCACTTCGAGGAGAAAATTGGCGCTCTGAAAGAGAACGTGAGCCACGTCGAATGACCT | 1365 |
| 2419 | Db | CAAGGCACTTCGAGGAGAAAATTGGCGCTCTGAAAGAGAACGTGAGCCACGTCGAATGACCT | 2478 |
| 1366 | QY | TGCTCGCCAGCTTACCACITTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACCTCTGGA | 1425 |
| 2479 | Db | TGCTCGCCAGCTTACCACITTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACCTCTGGA | 2538 |
| 1426 | QY | AGACCTGAACACCAGATGGAAGCTTCTGCAAGTGGCCGTCGAGGACCGAGTCAGGCAGCT | 1485 |
| 2539 | Db | AGACCTGAACACCAGATGGAAGCTTCTGCAAGTGGCCGTCGAGGACCGAGTCAGGCAGCT | 2598 |
| 1486 | QY | GCATGAAGCCACAGGGACTTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTGTCCA | 1545 |
| 2599 | Db | GCATGAAGCCACAGGGACTTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTGTCCA | 2658 |
| 1546 | QY | GGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCCAGAGAC | 1605 |
| 2659 | Db | GGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCCAGAGAC | 2718 |

1606 TCAACAACACTTGTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCT 1665
1666 GAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAAACCTCCGAAGACTGCAGAAAGGC 1725
1726 CCTTGTGCTTGGATCTCTTGAAGCTGTGAGCTGCTGATGATGCTTGGACCAAGCAACCT 1785
1786 CAAGCAAAATGACAGCCCATGGATATCCTGCGAGATTATTAATTGTTTGACCACTATTTA 1845
1846 TGACCGCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCTCTCTGCGTGGATATGTG 1905
1906 TCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCCTGTC 1965
1966 TTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGATACCT 2025
2026 TTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGATACCT 2085
2086 GCATGATTCTATCCAA 2101
3018 TGACCGCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCTCTCTGCGTGGATATGTG 3078
3079 TCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCCTGTC 3138
3139 TTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGATACCT 3198
3199 GCATGATTCTATCCAA 3214

RESULT 4

IS-09-845-416-10
Sequence 10, Application US/09845416
Publication No. US20030171312A1
GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE THEREOF
FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 3531
TYPE: DNA
ORGANISM: Homo sapiens
IS-09-845-416-10
Query Match 83.3%; Score 1750.8; DB 13; Length 3531;
Best Local Similarity 90.2%; Pred. No. 0;
Matches 1914; Conservative 0; Mismatches 187; Indels 21; Gaps 3;
1 GAGCTATGCCTACACAGAGCTGCTTATGTACACACCTCTGACCCCTACACGAGCCCAT 60
900 GAGCTATGCCTACACAGAGCTGCTTATGTACACACCTCTGACCCCTACACGAGCCCAT 959
61 TCCTTACAGCATTTGGAAGCTCCTGAAGACAAATGTCATTTGGCAGTTTATGATGGAGAG 120
960 TCCTTACAGCATTTGGAAGCTCCTGAAGACAAATGTCATTTGGCAGTTTATGATGGAGAG 1019
121 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAAGTATATCGTGGCTTCTTTC 180

1020 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1079
181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTTAATGATGTGGAAGTGTGAAAGA 240
1080 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTTAATGATGTGGAAGTGTGAAAGA 1139
241 CCAGTTTTCATCTCATGAGGGGTACATGATGGAATTGACAGCCCATCAAGGCCGGTGG 300
1140 CCAGTTTTCATCTCATGAGGGGTACATGATGGAATTGACAGCCCATCAAGGCCGGTGG 1199
301 TAATATTTCAATTTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
1200 TAATATTTCAATTTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 1259
361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGGATGGGAATGCCTCAGGGTAGC 420
1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGGATGGGAATGCCTCAGGGTAGC 1319
421 TAGCATGAAAAACAAAGCAATTTACATAGAGTT-----TTAATGGATCTCCAGAATCA 474
1320 TAGCATGAAAAACAAAGCAATTTACATAGAACTCATAGATTACTGCAACAGATTCCTCCCT 1379
475 GAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAACAAAGAAAAATGGA 534
1380 GGACCTGGAAAAAGTTTCTTGCCTGGCTTACAGAGCTGAAACAACTGCCAATGTCTTACA 1439
535 GGAAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACATTAAGGT 594
1440 GGATGCTACCCGTAAGGAAAGGCTCCTAGAAAGACTCCAAGGGAGTAAAAGAGCTGATGAA 1499
595 GCTTCAAGAAGATCTAGAACAAGAACAAAGTCAAGGTCAATTTCTCTCACTCACATGTTGT 654
1500 ACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCACAACCTGGA 1559
655 GGTAG-----TTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGA 702
1560 TGAACACAGCCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGAGTCTCTGTTACA 1619
703 ACAACTTAAGG---TATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGCCG 759
1620 AAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACAT 1679
760 CTGGGTTCTTTTACAAGACAGTTCTGACCCAGTGAAGCGTCTGACCTTTCTCTGCAGGA 819
1680 TAGGTCCCATTTGGAAGCCAGTTCTGACCCAGTGAAGCGTCTGACCTTTCTCTGCAGGA 1739
820 ACTTCTGGTGTGGTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGG 879
1740 ACTTCTGGTGTGGTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGG 1799
880 CGACTTTCCAGCAGTTTCAAGAGCAGAACCGATGATAGGCGCTTCAAGAGGGAATTGAA 939
1800 CGACTTTCCAGCAGTTTCAAGAGCAGAACCGATGATAGGCGCTTCAAGAGGGAATTGAA 1859
940 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTTGACAGAGCA 999
1860 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTTGACAGAGCA 1919
1000 GCCTTTGGAAGGACTAGAGAAACTCTACCCAGAGCCCGAGAGGCTGCCTCCTGAGGAGAG 1059
1920 GCCTTTGGAAGGACTAGAGAAACTCTACCCAGAGCCCGAGAGGCTGCCTCCTGAGGAGAG 1979
1060 AGCCCAGAATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATAGTGGTGGGA 1119
1980 AGCCCAGAATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATAGTGGTGGGA 2039
1120 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACT 1179
2040 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACT 2099
1180 CCAGGAACCTTCAAGAGCCACCGATGAGCTGGAACCTCAAGCTGCGCAAGCTGAGGTGAT 1239
2100 CCAGGAACCTTCAAGAGCCACCGATGAGCTGGAACCTCAAGCTGCGCAAGCTGAGGTGAT 2159


```
QY 1240 CAAGGATCCTGGCAGCCCGTGGCGGATCTCCTCAATGACTCTCTCCAAAGATCACCTCGA 1299
    |||
Db 2160 CAAGGATCCTGGCAGCCCGTGGCGGATCTCCTCAATGACTCTCTCCAAAGATCACCTCGA 2219
    |||
QY 1300 GAAAGTCAAGGCACCTTCGAGGAGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAA 1359
    |||
Db 2220 GAAAGTCAAGGCACCTTCGAGGAGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAA 2279
    |||
QY 1360 TGACCTTGCTCGCCAGCTTACACATTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 1419
    |||
Db 2280 TGACCTTGCTCGCCAGCTTACACATTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 2339
    |||
QY 1420 TCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAAGTGGCCGTCGAGGACCGAGTCAG 1479
    |||
Db 2340 TCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAAGTGGCCGTCGAGGACCGAGTCAG 2399
    |||
QY 1480 GCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTC 1539
    |||
Db 2400 GCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTC 2459
    |||
QY 1540 TGTCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 1599
    |||
Db 2460 TGTCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 2519
    |||
QY 1600 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTTAGC 1659
    |||
Db 2520 CGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTTAGC 2579
    |||
QY 1660 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 1719
    |||
Db 2580 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 2639
    |||
QY 1720 GAAGGCCCTTTGCTTGGATCTCTGAGCCTGTCTAGCTGCATGTGATGCCCTTGACCCAGCA 1779
    |||
Db 2640 GAAGGCCCTTTGCTTGGATCTCTGAGCCTGTCTAGCTGCATGTGATGCCCTTGACCCAGCA 2699
    |||
QY 1780 CAACCTCAAGCAAAATGACCGCCATGGATATCTCTCAGATTATTAATTGTTGACCCAC 1839
    |||
Db 2700 CAACCTCAAGCAAAATGACCGCCATGGATATCTCTCAGATTATTAATTGTTGACCCAC 2759
    |||
QY 1840 TATTTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCCTCTCTGCGTGGA 1899
    |||
Db 2760 TATTTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCCTCTCTGCGTGGA 2819
    |||
QY 1900 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1959
    |||
Db 2820 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 2879
    |||
QY 1960 CCTGTCTTTAAAACTGGCATCAATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 2019
    |||
Db 2880 CCTGTCTTTAAAACTGGCATCAATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 2939
    |||
QY 2020 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 2079
    |||
Db 2940 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 2999
    |||
QY 2080 CCTTCTGCATGATTCTATCCAA 2101
    |||
Db 3000 CCTTCTGCATGATTCTATCCAA 3021
    |||
```

RESULT 5
US-09-845-416-30
; Sequence 30, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30

```
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 4498
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-30

Query Match      83.3%; Score 1750.8; DB 13; Length 4498;
Best Local Similarity 90.2%; Pred. No. 0;
Matches 1914; Conservative 0; Mismatches 187; Indels 21; Gaps 3;

QY 1 GAGTATGCCTACACACAGGCTGCTTATGTCTACCCACCTCTGACCCCTACACGAGCCCAT 60
    |||
Db 1657 GAGTATGCCTACACACAGGCTGCTTATGTCTACCCACCTCTGACCCCTACACGAGCCCAT 1716
    |||
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 120
    |||
Db 1717 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG 1776
    |||
QY 121 TGAAGTAAACCTGGACCGTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 180
    |||
Db 1777 TGAAGTAAACCTGGACCGTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 1836
    |||
QY 181 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 240
    |||
Db 1837 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTCTTAATGATGTGGAAGTGGTGAAGA 1896
    |||
QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGTTGG 300
    |||
Db 1897 CCAGTTTCATCTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGTTGG 1956
    |||
QY 301 TAATATTCTACAATTGGGAAGTAAAGTCTCTCTAAATCAAGATGGGAATGCCTCAGGGTAGC 420
    |||
Db 1957 TAATATTCTACAATTGGGAAGTAAAGTCTCTCTAAATCAAGATGGGAATGCCTCAGGGTAGC 2016
    |||
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATCAAGATGGGAATGCCTCAGGGTAGC 420
    |||
Db 2017 AACTGAAGTACAAGAGCAGATGAATCTCTCTAAATCAAGATGGGAATGCCTCAGGGTAGC 2076
    |||
QY 421 TAGCATGAAAAACAAGCAATTTACATAGAGTT-----TTAATGGATCTCCAGAATCA 474
    |||
Db 2077 TAGCATGAAAAACAAGCAATTTACATAGAACTCATAGATTACTGCAACAGTTCCCTCT 2136
    |||
QY 475 GAACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAAGAAACAAAGGAAAAATGGA 534
    |||
Db 2137 GGACCTGGAAGAGTTCTTGCTGGCTTACAGAAAGCTGAAACAACTGCCAATGTCTCTACA 2196
    |||
QY 535 GGAAGAGCCTCTTGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACACATAAGGT 594
    |||
Db 2197 GGATGCTACCCGTAAGGAAAGCTCCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAA 2256
    |||
QY 595 GCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATCTCTCACTCACATGGTGT 654
    |||
Db 2257 ACAATGCAAGACCTCCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACTGGA 2316
    |||
QY 655 GGTAG-----TTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGGAAGA 702
    |||
Db 2317 TGAACACAGCCAAATAATCCTGAGATCCCTGGAGGTTCCGATGATGCACTCCTGTTACA 2376
    |||
QY 703 ACAACTTAAGG---TATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCG 759
    |||
Db 2377 AAGACGTTTGGATAACATGAACCTCAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACAT 2436
    |||
QY 760 CTGGGTCTTTTACAAGACAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGA 819
    |||
Db 2437 TAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGA 2496
    |||
QY 820 ACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAGCCGGCAGGCACCTATTGGAGG 879
    |||
Db 2497 ACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAGCCGGCAGGCACCTATTGGAGG 2556
    |||
```

QY 880 CGACTTTCCAGCAGTTTCAAGACGACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA 939
Db 2557 CGACTTTCCAGCAGTTTCAAGACGACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA 2616
QY 940 AACTAAAGAACTTGAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA 999
Db 2617 AACTAAAGAACTTGAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA 2676
QY 1000 GCCTTTGGAAGGACTAGAGAACTCTACCAAGGAGCCCGAGAGCTGCCTCCTGAGGAGAG 1059
Db 2677 GCCTTTGGAAGGACTAGAGAACTCTACCAAGGAGCCCGAGAGCTGCCTCCTGAGGAGAG 2736
QY 1060 AGCCCAGAATGTCTCGGCTTCTACGAAGACGAGCTGAGGAGGTCAATACTAGTGGGA 1119
Db 2737 AGCCCAGAATGTCTCGGCTTCTACGAAGACGAGCTGAGGAGGTCAATACTAGTGGGA 2796
QY 1120 AAAATTGAACCTTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 1179
Db 2797 AAAATTGAACCTTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 2856
QY 1180 CCAGGAACCTTCAAGAGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 1239
Db 2857 CCAGGAACCTTCAAGAGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 2916
QY 1240 CAAGGGATCCTGGCAGCCCGTGGGGGATCTCCTCATTTGACTCTCTCCTCAAGATCACCTCGA 1299
Db 2917 CAAGGGATCCTGGCAGCCCGTGGGGGATCTCCTCATTTGACTCTCTCCTCAAGATCACCTCGA 2976
QY 1300 GAAAGTCAAGGCACTTCGAGGAGAAATTCGGCCCTCTGAAAGAGAACGTTGAGCCACGTCAA 1359
Db 2977 GAAAGTCAAGGCACTTCGAGGAGAAATTCGGCCCTCTGAAAGAGAACGTTGAGCCACGTCAA 3036
QY 1360 TGACCTTGCTCGCAGCTTACCACCTTTGGGCATTCAGCTCTCAACCGTATAACCTCAGCAC 1419
Db 3037 TGACCTTGCTCGCAGCTTACCACCTTTGGGCATTCAGCTCTCAACCGTATAACCTCAGCAC 3096
QY 1420 TCTGGAAGACCTGAAACACCAAGATGGAAGCTTCTGCAGGTGGCCCTCGAGGACCGAGTCAG 1479
Db 3097 TCTGGAAGACCTGAAACACCAAGATGGAAGCTTCTGCAGGTGGCCCTCGAGGACCGAGTCAG 3156
QY 1480 GCAGCTGCATGAAGCCACACAGGACTTTGGTCCAGCATCTCAGACATTTCTTCCAGTC 1539
Db 3157 GCAGCTGCATGAAGCCACACAGGACTTTGGTCCAGCATCTCAGACATTTCTTCCAGTC 3216
QY 1540 TGTCCAGGTCCTTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACCA 1599
Db 3217 TGTCCAGGTCCTTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACCA 3276
QY 1600 CGAGACTCAAAACAACTTGTGGGACCATCCCAAATGACAGAGCTCTACCAGTCTTTAGC 1659
Db 3277 CGAGACTCAAAACAACTTGTGGGACCATCCCAAATGACAGAGCTCTACCAGTCTTTAGC 3336
QY 1660 TGACCTGAATAATGTGAGTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 1719
Db 3337 TGACCTGAATAATGTGAGTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3396
QY 1720 GAAGGCCCTTGTCTTGGATCTCTTGAGCCTGTGAGCTGCAATGATGCCTTTGGACAGCA 1779
Db 3397 GAAGGCCCTTGTCTTGGATCTCTTGAGCCTGTGAGCTGCAATGATGCCTTTGGACAGCA 3456
QY 1780 CAACCTCAAGCAAAATGACCAAGCCATGGATATCCTGCAGATTAATAATTGTTTGAACCA 1839
Db 3457 CAACCTCAAGCAAAATGACCAAGCCATGGATATCCTGCAGATTAATAATTGTTTGAACCA 3516
QY 1840 TATTTATGACCGCCTGGAGCAAGAGCAACAAATTTGGTCAACGTCCCTCTCTGCGTGA 1899
Db 3517 TATTTATGACCGCCTGGAGCAAGAGCAACAAATTTGGTCAACGTCCCTCTCTGCGTGA 3576
QY 1900 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1959
Db 3577 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 3636

QY 1960 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 2019
Db 3637 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 3696
QY 2020 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGGCCT 2079
Db 3697 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGGCCT 3756
QY 2080 CCTTCTGCATGATTTCTATCCAA 2101
Db 3757 CCTTCTGCATGATTTCTATCCAA 3778
RESULT 6
US-09-845-416-9
; Sequence 9, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3858
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-9
Query Match 83.0%; Score 1743; DB 13; Length 3858;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 348; Gaps 1;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACAGCTTTAGAAAGATTAATCGTGGCTTCTTC 60
Db 900 GAGCTATGCCTACACACAGGCTGCTTATGTCAACACCTCTGACCTACACGGAGCCCAT 959
QY 61 TCCTTACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCAATGATGGAGAG 120
Db 960 TCCTTACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCAATGATGGAGAG 1019
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATTAATCGTGGCTTCTTC 180
Db 1020 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGATTAATCGTGGCTTCTTC 1079
QY 181 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1080 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1139
QY 241 CCAGTTTCATCTACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 300
Db 1140 CCAGTTTCATCTACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGG 1199
QY 301 TAATATTCTACAATTGGGAAGTAACTGATTGGAAACAGGAAAATTATCAGAAGATGAAGA 360
Db 1200 TAATATTCTACAATTGGGAAGTAACTGATTGGAAACAGGAAAATTATCAGAAGATGAAGA 1259
QY 361 AACTGAAGTACAAGAGCAGATGATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
Db 1260 AACTGAAGTACAAGAGCAGATGATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1319
QY 421 TAGCATGGAATAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAACTCAGAACT 480
Db 1320 TAGCATGGAATAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAACTCAGAACT 1379
QY 481 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAAGAAATGGAGGAAGA 540
Db 1380 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAAGAAATGGAGGAAGA 1439

QY 541 GCCTCTGGACCTGATCTTGAAGACCTAAACGCCAAGTACAAACAACATAAGGTGCTTCA 600
Db 1440 GCCTCTGGACCTGATCTTGAAGACCTAAACGCCAAGTACAAACAACATAAGGTGCTTCA 1499
QY 601 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGTTGGTAGT 660
Db 1500 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCTCACTCACATGTTGGTAGT 1559
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGTTTGAAGAACAACTTAAGGTATTGGG 720
Db 1560 TGATGAATCTAGTGGAGATCACGCAACTGCTGTTTGAAGAACAACTTAAGGTATTGGG 1619
QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGTTCCTTTACAAGA --- 777
Db 1620 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGTTCCTTTACAAGACAC 1679
QY 778 ----- 777
Db 1680 TCATAGATTACTGCAACAGTTCCCTGGACCTGGAAAAGTTTCTTGGCTGCTTACAGA 1739
QY 778 ----- 777
Db 1740 AGCTGAAAACAACTGCCAATGTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGA 1799
QY 778 ----- 777
Db 1800 CTCCAAGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGAAAATTGAAGC 1859
QY 778 ----- 777
Db 1860 TCACACAGATGTTTATCAACAACCTGGATGAACAGCCAAAAAATCCTGAGATCCCTGGA 1919
QY 778 ----- 777
Db 1920 AGGTTCCGATGATGCAGTCTCTTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAG 1979
QY 778 ----- CAGTTCTGACCAGTG 792
Db 1980 TGAACCTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTG 2039
QY 793 GAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGA 852
Db 2040 GAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGA 2099
QY 853 ATTAAGCCGGCAGGCACCTATTGGAGGCGCACTTTCAGCAGTTTCAGAAGCAGAACGATGT 912
Db 2100 ATTAAGCCGGCAGGCACCTATTGGAGGCGCACTTTCAGCAGTTTCAGAAGCAGAACGATGT 2159
QY 913 ACATAGGCGCTTCAAGAGGGAAATTGAAAACCTAAGAACTGTAATCATGAGTACTCTTTGA 972
Db 2160 ACATAGGCGCTTCAAGAGGGAAATTGAAAACCTAAGAACTGTAATCATGAGTACTCTTTGA 2219
QY 973 GACTGTAGCAATATTTCTGACAGAGCAGCCTTTTGGAAAGGACTAGAGAAAACCTTACCAGGA 1032
Db 2220 GACTGTAGCAATATTTCTGACAGAGCAGCCTTTTGGAAAGGACTAGAGAAAACCTTACCAGGA 2279
QY 1033 GCCCAGAGAGCTGCCCTCCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTAGCAAAAGCA 1092
Db 2280 GCCCAGAGAGCTGCCCTCCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTAGCAAAAGCA 2339
QY 1093 GGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 1152
Db 2340 GGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 2399
QY 1153 AAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGA 1212
Db 2400 AAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGA 2459
QY 1213 CCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTTGGCAGCCCGTGGGCGATCTCCT 1272
Db 2460 CCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTTGGCAGCCCGTGGGCGATCTCCT 2519

QY 1273 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGC 1332
Db 2520 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGC 2579
QY 1333 TCTGAAAAGAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAT 1392
Db 2580 TCTGAAAAGAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAT 2639
QY 1393 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAAGTGAAGCTTCT 1452
Db 2640 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAAGTGAAGCTTCT 2699
QY 1453 GCAGGTGGCCGTCGAGGACCGAGTCAGGCGAGCTGCAATGAAGCCACACAGGACTTTGGTCC 1512
Db 2700 GCAGGTGGCCGTCGAGGACCGAGTCAGGCGAGCTGCAATGAAGCCACACAGGACTTTGGTCC 2759
QY 1513 AGCATCTCAGCACCTTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 1572
Db 2760 AGCATCTCAGCACCTTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 2819
QY 1573 AAACAAAAGTCCCTACTATATCAACCCAGAGACTCAAAACAACTTGTGGGACCATCCCAA 1632
Db 2820 AAACAAAAGTCCCTACTATATCAACCCAGAGACTCAAAACAACTTGTGGGACCATCCCAA 2879
QY 1633 AATGACAGAGCTCTACCACTTTTAGCTGACCTGAAATAATGTGAGATTCAGCTTATAG 1692
Db 2880 AATGACAGAGCTCTACCACTTTTAGCTGACCTGAAATAATGTGAGATTCAGCTTATAG 2939
QY 1693 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 1752
Db 2940 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 2999
QY 1753 AGCTGCATGTGATGCCCTTGGACCAAGCAACCTCAAGCAAAATGACCCAGCCCATGGATAT 1812
Db 3000 AGCTGCATGTGATGCCCTTGGACCAAGCAACCTCAAGCAAAATGACCCAGCCCATGGATAT 3059
QY 1813 CCTGCAGATTATTAATTGTTGACCACTATTTATGACCGCCTGGAGCAAGAGCAACAACAA 1872
Db 3060 CCTGCAGATTATTAATTGTTGACCACTATTTATGACCGCCTGGAGCAAGAGCAACAACAA 3119
QY 1873 TTTGGTCAACGTCCTCTCTGCTGGATATGTTGTAACCTGAACTGGCTGCTGAATGTTTATGA 1932
Db 3120 TTTGGTCAACGTCCTCTCTGCTGGATATGTTGTAACCTGAACTGGCTGCTGAATGTTTATGA 3179
QY 1933 TACGGACGAACAGGAGGATCCGTGCTCTTTTAAACTGGCATCATTTCCCTGTG 1992
Db 3180 TACGGACGAACAGGAGGATCCGTGCTCTTTTAAACTGGCATCATTTCCCTGTG 3239
QY 1993 TAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 2052
Db 3240 TAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 3299
QY 2053 ATTTTGTGACCAAGCGCAGGCTGGGCTCCCTTCTGCAATGATTTATCCAA 2101
Db 3300 ATTTTGTGACCAAGCGCAGGCTGGGCTCCCTTCTGCAATGATTTATCCAA 3348

RESULT 7

US-09-845-416-29
; Sequence 29, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29

| | | | | | |
|----|------|--|------|--------------------------------------|--|
| | | ; LENGTH: 4825 | | | |
| | | ; TYPE: DNA | | | |
| | | ; ORGANISM: Homo sapiens | | | |
| | | US-09-845-416-29 | | | |
| | | Query Match | | Score 1743; DB 13; Length 4825; | |
| | | Best Local Similarity | | 85.8%; Pred. No. 0; | |
| | | Matches 2101; Conservative | | 0; Mismatches 0; Indels 348; Gaps 1; | |
| QY | 1 | GAGCTATGCCTACACACAGGCTGCTTATGTACCAACCTCTGACCCCTACACGGAGCCCAT | 60 | | |
| Db | 1657 | GAGCTATGCCTACACACAGGCTGCTTATGTACCAACCTCTGACCCCTACACGGAGCCCAT | 1716 | | |
| QY | 61 | TCCTTCACAGCATTTGGAAGCTCTGGAAGACAAGTCAATTTGGCAGTTCAATGATGGAGAG | 120 | | |
| Db | 1717 | TCCTTCACAGCATTTGGAAGCTCTGGAAGACAAGTCAATTTGGCAGTTCAATGATGGAGAG | 1776 | | |
| QY | 121 | TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC | 180 | | |
| Db | 1777 | TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC | 1836 | | |
| QY | 181 | TGCTGAGGACACATTTGCAAGCAAGGAGAGATTTCTTAATGATGTGGAAGTGTGAAAGA | 240 | | |
| Db | 1837 | TGCTGAGGACACATTTGCAAGCAAGGAGAGATTTCTTAATGATGTGGAAGTGTGAAAGA | 1896 | | |
| QY | 241 | CCAGTTTCATACTCATGAGGGTACATGATGGATTTGACAGCCCATCAGGCGGGTTGG | 300 | | |
| Db | 1897 | CCAGTTTCATACTCATGAGGGTACATGATGGATTTGACAGCCCATCAGGCGGGTTGG | 1956 | | |
| QY | 301 | TAATATTCTACAATTTGGGAAGTAAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA | 360 | | |
| Db | 1957 | TAATATTCTACAATTTGGGAAGTAAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA | 2016 | | |
| QY | 361 | AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC | 420 | | |
| Db | 2017 | AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC | 2076 | | |
| QY | 421 | TAGCATGGAATAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT | 480 | | |
| Db | 2077 | TAGCATGGAATAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT | 2136 | | |
| QY | 481 | GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAGAAAGAACAGGAAAAATGGAGGAAGA | 540 | | |
| Db | 2137 | GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAGAAAGAACAGGAAAAATGGAGGAAGA | 2196 | | |
| QY | 541 | GCCTCTTGACCTGATCTTTGAAGACCTTAAACCGCCAAAGTACAACACATAGGTGCTTCA | 600 | | |
| Db | 2197 | GCCTCTTGACCTGATCTTTGAAGACCTTAAACCGCCAAAGTACAACACATAGGTGCTTCA | 2256 | | |
| QY | 601 | AGAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTTCTCACTCACATGGTGGTGTAGT | 660 | | |
| Db | 2257 | AGAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTTCTCACTCACATGGTGGTGTAGT | 2316 | | |
| QY | 661 | TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG | 720 | | |
| Db | 2317 | TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATTGGG | 2376 | | |
| QY | 721 | AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGA--- | 777 | | |
| Db | 2377 | AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAC | 2436 | | |
| QY | 778 | ----- | 777 | | |
| Db | 2437 | TCATAGATTACTGCAACAGTTCCCTCGACCTTGAAAGTTTCTTGCTGGCTTACAGA | 2496 | | |
| QY | 778 | ----- | 777 | | |
| Db | 2497 | AGCTGAACAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGGAAGGCTCCTAGAAGA | 2556 | | |
| QY | 778 | ----- | 777 | | |
| Db | 2557 | CTCCAAGGGAGTAAAGAGCTGATGAACAAATGGCAAGACCTCCAAGGTGAATTAAGAC | 2616 | | |

| | | | |
|----|------|---|------|
| QY | 778 | ----- | 777 |
| Db | 2617 | TCACACAGATGTTTATCACAACTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGA | 2676 |
| QY | 778 | ----- | 777 |
| Db | 2677 | AGGTTCCGATGATGCAGTCCCTGTTTACAAAAGACGTTTGGATPAACATGAACCTTCAAGTGGAG | 2736 |
| QY | 778 | ----- | 792 |
| Db | 2737 | TGAACCTCGAAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTG | 2796 |
| QY | 793 | GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTTGGTGTGGCTACAGCTGAAAAGATGATGA | 852 |
| Db | 2797 | GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTTGGTGTGGCTACAGCTGAAAAGATGATGA | 2856 |
| QY | 853 | ATTAAAGCGGCGAGGCACCTATTGGAGGCGACTTTCAGCAGTTTCAGAAAGCAGAACGATGT | 912 |
| Db | 2857 | ATTAAAGCGGCGAGGCACCTATTGGAGGCGACTTTCAGCAGTTTCAGAAAGCAGAACGATGT | 2916 |
| QY | 913 | ACATAGGCGCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGA | 972 |
| Db | 2917 | ACATAGGCGCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGA | 2976 |
| QY | 973 | GACTGTACGAATATTTCTGCAGAGCAGCCCTTTTGGAGGACTAGAGAAACTCTACCAGGA | 1032 |
| Db | 2977 | GACTGTACGAATATTTCTGCAGAGCAGCCCTTTTGGAGGACTAGAGAAACTCTACCAGGA | 3036 |
| QY | 1033 | GCCACAGAGCTGCCCTCCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCA | 1092 |
| Db | 3037 | GCCACAGAGCTGCCCTCCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCA | 3096 |
| QY | 1093 | GGCTGAGGAGGTCAATATCTGAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAG | 1152 |
| Db | 3097 | GGCTGAGGAGGTCAATATCTGAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAG | 3156 |
| QY | 1153 | AAAAATAGATGAGACCCCTTGAAGAAGTCCAGGAACCTCAAGAGGCCACCGATGAGCTGGA | 1212 |
| Db | 3157 | AAAAATAGATGAGACCCCTTGAAGAAGTCCAGGAACCTCAAGAGGCCACCGATGAGCTGGA | 3216 |
| QY | 1213 | CCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGCATCTCCT | 1272 |
| Db | 3217 | CCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGCATCTCCT | 3276 |
| QY | 1273 | CATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGAAAATTGCGCC | 1332 |
| Db | 3277 | CATTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGAAAATTGCGCC | 3336 |
| QY | 1333 | TCTGAAAAGAGAACGTCAGCCACGTCATGACCTTGTCTGCGCCAGCTTACCACCTTTGGGCAT | 1392 |
| Db | 3337 | TCTGAAAAGAGAACGTCAGCCACGTCATGACCTTGTCTGCGCCAGCTTACCACCTTTGGGCAT | 3396 |
| QY | 1393 | TCAGTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCT | 1452 |
| Db | 3397 | TCAGTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCT | 3456 |
| QY | 1453 | GCAGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGGACTTTGGTCC | 1512 |
| Db | 3457 | GCAGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGGACTTTGGTCC | 3516 |
| QY | 1513 | AGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC | 1572 |
| Db | 3517 | AGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC | 3576 |
| QY | 1573 | AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACAACTTGTCTGGGACCATCCCAA | 1632 |
| Db | 3577 | AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACAACTTGTCTGGGACCATCCCAA | 3636 |
| QY | 1633 | AATGACAGAGCTCTACCAAGTCTTTAGCTGACCTGAATAATGTGATTTCTCAGCTTATAG | 1692 |
| Db | 3637 | AATGACAGAGCTCTACCAAGTCTTTAGCTGACCTGAATAATGTGATTTCTCAGCTTATAG | 3696 |
| QY | 1693 | GACTGCCATGAAAACCTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC | 1752 |

|||||
Db 3697 GACTGCCATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 3756
|||||
QY 1753 AGCTGATGTGATGCCCTTGGACGAGCAAACTCAAGCAAAATGACCGCCATGGATAT 1812
|||||
Db 3757 AGCTGATGTGATGCCCTTGGACGAGCAAACTCAAGCAAAATGACCGCCATGGATAT 3816
|||||
QY 1813 CCTGCAGATTATTAAATGTTTGACCACTATTTATGACCGCCTGGAGCAAGACAAACAA 1872
|||||
Db 3817 CCTGCAGATTATTAAATGTTTGACCACTATTTATGACCGCCTGGAGCAAGACAAACAA 3876
|||||
QY 1873 TTTGGTCAACGTCCTCTCTCGGTGGATATGTGCTGAACTGGCTGCTGAATGTTTATGA 1932
|||||
Db 3877 TTTGGTCAACGTCCTCTCTCGGTGGATATGTGCTGAACTGGCTGCTGAATGTTTATGA 3936
|||||
QY 1933 TACGGGACGAACAGGAGGATCCGTGTCCTGTCTTTTAAAACTGGCATCAATTCCTGTG 1992
|||||
Db 3937 TACGGGACGAACAGGAGGATCCGTGTCCTGTCTTTTAAAACTGGCATCAATTCCTGTG 3996
|||||
QY 1993 TAAAGCACATTGGAAGACAAAGTACAGATACCTTTTCAAGCAAAGTGGCAAGTTCAACAGG 2052
Db 3997 TAAAGCACATTGGAAGACAAAGTACAGATACCTTTTCAAGCAAAGTGGCAAGTTCAACAGG 4056
|||||
QY 2053 ATTTTGTGACGCGCAGGCTGGGCCTCCTTCTGATGATTTCTATCCAA 2101
|||||
Db 4057 ATTTTGTGACGCGCAGGCTGGGCCTCCTTCTGATGATTTCTATCCAA 4105
|||||

RESULT 8

US-09-845-416-35
; Sequence 35, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 4848
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-35

Query Match 83.0%; Score 1743; DB 13; Length 4848;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 348; Gaps 1;
QY 1 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACCGGAGCCCAT 60
Db 1680 GAGCTATGCCCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCCTACACCGGAGCCCAT 1739
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
Db 1740 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1799
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 180
Db 1800 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 1859
QY 181 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1860 TGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1919
QY 241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTTGACGCCCATCAGGGCCGGTTGG 300
Db 1920 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTTGACGCCCATCAGGGCCGGTTGG 1979
|||||

QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAATGAAGA 360
Db 1980 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAATGAAGA 2039
|||||
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCCTCAGGGTAGC 420
Db 2040 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCCCTCAGGGTAGC 2099
|||||
QY 421 TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 480
Db 2100 TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 2159
|||||
QY 481 GAAAGAGTTGAATGACTGCTAAACAAACAAAGAAAGAAACAAAGAAATGGAGGAAGA 540
Db 2160 GAAAGAGTTGAATGACTGCTAAACAAACAAAGAAAGAAACAAAGAAATGGAGGAAGA 2219
|||||
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACAACATAAGGTGCTTCA 600
Db 2220 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACAACATAAGGTGCTTCA 2279
|||||
QY 601 AGAAGATCTAGAACAAAGAACAGTCAGGGTCAATTTCTCACTCACATGGTGGTGTAGT 660
Db 2280 AGAAGATCTAGAACAAAGAACAGTCAGGGTCAATTTCTCACTCACATGGTGGTGTAGT 2339
|||||
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAAACAACCTTAAGGTATTGGG 720
Db 2340 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAAACAACCTTAAGGTATTGGG 2399
|||||
QY 721 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGTTCTTTTACAAGA --- 777
Db 2400 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGTTCTTTTACAAGACAC 2459
|||||
QY 778 --- --- --- 777
Db 2460 TCATAGATTACTGCAACAGTCCCCCTGGACCTGGAAAAAGTTTCTTGGCTGGCTTACAGA 2519
|||||
QY 778 --- --- --- 777
Db 2520 AGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGTAAGAAAGGCTCCTAGAAGA 2579
|||||
QY 778 --- --- --- 777
Db 2580 CTCCAAGGGAGTAAAAGAGCTGATGAAACAAATGGCAAGACCTCCAAGGTGAAATTTGAAGC 2639
|||||
QY 778 --- --- --- 777
Db 2640 TCACACAGATGTTTATCACAACCTGGATGAAACAGCAAGCAAGCAAGCAAGCAAGCAAGCA 2699
|||||
QY 778 --- --- --- 777
Db 2700 AGGTCCGATGATGCAGTCCCTGTATCAAAAGACGTTTGGATAACATGAACATTTCAAGTGGAG 2759
|||||
QY 778 --- --- --- 778
Db 2760 TGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCCAGTG 2819
|||||
QY 793 GAAGCTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGGCTACAGCTGAAAGATGATGA 852
Db 2820 GAAGCTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGGCTACAGCTGAAAGATGATGA 2879
|||||
QY 853 ATTAAGCCGGCAGGCACCTATTGGAGGGCAGCTTTCCAGCAGTTTCAAGAGCAGAACGATGT 912
Db 2880 ATTAAGCCGGCAGGCACCTATTGGAGGGCAGCTTTCCAGCAGTTTCAAGAGCAGAACGATGT 2939
|||||
QY 913 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGA 972
Db 2940 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCATGAGTACTCTTGA 2999
|||||
QY 973 GACTGTACGAATATTTCTGACAGAGCAGCCCTTTTGAAGGACTAGAGAAACTCTACCAGGA 1032
Db 3000 GACTGTACGAATATTTCTGACAGAGCAGCCCTTTTGAAGGACTAGAGAAACTCTACCAGGA 3059
|||||
QY 1033 GCCCAGAGAGCTGCCTCCTTGAGGAGAGAGGCCAGAAATGTCACTCGGCTTCTACGAAAGCA 1092
|||||

Db 3060 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGCTTCTACGAAAGCA 3119
QY 1093 GGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAG 1152
Db 3120 GGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAG 3179
QY 1153 AAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGA 1212
Db 3180 AAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGA 3239
QY 1213 CCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGSCAGCCCCGTGGCGGATCTCCT 1272
Db 3240 CCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGSCAGCCCCGTGGCGGATCTCCT 3299
QY 1273 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCC 1332
Db 3300 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCC 3359
QY 1333 TCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAT 1392
Db 3360 TCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAT 3419
QY 1393 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAGATGGAAGCTTCT 1452
Db 3420 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAGATGGAAGCTTCT 3479
QY 1453 GCAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTGGTCC 1512
Db 3480 GCAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTGGTCC 3539
QY 1513 AGCATCTCAGCACTTTCTTTCCAGCTGTGCCAGGTCCCTGGGAGAGAGCCATCTCGCC 1572
Db 3540 AGCATCTCAGCACTTTCTTTCCAGCTGTGCCAGGTCCCTGGGAGAGAGCCATCTCGCC 3599
QY 1573 AAACAAAGTGCCCTACTATATCAACACAGAGACTCAAAACAACCTTGCTGGGACCATCCCAA 1632
Db 3600 AAACAAAGTGCCCTACTATATCAACACAGAGACTCAAAACAACCTTGCTGGGACCATCCAA 3659
QY 1633 AATGACAGAGCTCTACCACTCTTTAGCTGACCTGAATAATGTCAAGATTCTCAGCTTATAG 1692
Db 3660 AATGACAGAGCTCTACCACTCTTTAGCTGACCTGAATAATGTCAAGATTCTCAGCTTATAG 3719
QY 1693 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 1752
Db 3720 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 3779
QY 1753 AGCTGCATGTGATGCCCTTGGACCGACACAACTCAAGCAAAATGACCAGCCCATGGATAT 1812
Db 3780 AGCTGCATGTGATGCCCTTGGACCGACACAACTCAAGCAAAATGACCAGCCCATGGATAT 3839
QY 1813 CCTGCAGATTATTAATTGTTTGACCACTATTTATGACCCGCTGGAGCAAGAGCAACAA 1872
Db 3840 CCTGCAGATTATTAATTGTTTGACCACTATTTATGACCCGCTGGAGCAAGAGCAACAA 3899
QY 1873 TTTGGTCAACGTCCCTCTCTGCGTGGATATGTGTGAATGGCTGCTGAATGTTTATGA 1932
Db 3900 TTTGGTCAACGTCCCTCTCTGCGTGGATATGTGTGAATGGCTGCTGAATGTTTATGA 3959
QY 1933 TACGGGACGAACAGGGAGGATCCGTGTCCTGTCTTTTAAACTGGCATCATTTCCCTGTG 1992
Db 3960 TACGGGACGAACAGGGAGGATCCGTGTCCTGTCTTTTAAACTGGCATCATTTCCCTGTG 4019
QY 1993 TAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 2052
Db 4020 TAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 4079
QY 2053 ATTTTGTGACAGCGCAGGCTGGGCTCCTTCTGTCATGATTCTATCCAA 2101
Db 4080 ATTTTGTGACAGCGCAGGCTGGGCTCCTTCTGTCATGATTCTATCCAA 4128

US-09-845-416-36
; Sequence 36, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 5060
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-36
Query Match 83.0%; Score 1743; DB 13; Length 5060;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 348; Gaps 1;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGGAGCCCAT 60
Db 1892 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGGAGCCCAT 1951
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCAATTGATGGAGAG 120
Db 1952 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCAATTGATGGAGAG 2011
QY 121 TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
Db 2012 TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTC 2071
QY 181 TGCTGAGGACACATTGCAAGCAAGGAGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 2072 TGCTGAGGACACATTGCAAGCAAGGAGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 2131
QY 241 CCAGTTTCATACCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG 300
Db 2132 CCAGTTTCATACCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG 2191
QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 360
Db 2192 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 2251
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGTTAGC 420
Db 2252 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGTTAGC 2311
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480
Db 2312 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 2371
QY 481 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAGAAAGAAACAAGGAAATGGAGGAAGA 540
Db 2372 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAGAAAGAAACAAGGAAATGGAGGAAGA 2431
QY 541 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAAACAATAGGTGCTTCA 600
Db 2432 GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAAACAATAGGTGCTTCA 2491
QY 601 AGAAGATCTAGAACCAAGAACCAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 660
Db 2492 AGAAGATCTAGAACCAAGAACCAAGTCAGGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 2551
QY 661 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATGGG 720
Db 2552 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGGAAAGAACAACTTAAGGTATGGG 2611
QY 721 AGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGTCTTTTACAGA --- 777

Db 2612 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAC 2671
QY 778 ----- 777
Db 2672 TCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGGCTGGCTTACAGA 2731
QY 778 ----- 777
Db 2732 AGCTGAAACAACTGCAATGTCTACAGGATGCTACCCGTAAAGGAAAGGCTCTCTAGAAAGA 2791
QY 778 ----- 777
Db 2792 CTCCAAGGAGTAAAGAGCTGTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 2851
QY 778 ----- 777
Db 2852 TCACACAGATGTTTATCAACAACCTGGATGAAACAGCCAAAAAATCCTGAGATCCCTGGA 2911
QY 778 ----- 777
Db 2912 AGGTTCCGATGATGCAGTCCCTGTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAG 2971
QY 778 ----- CAGTTCTGACCCAGTG 792
Db 2972 TGAACCTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAAGCCAGTTCTGACCCAGTG 3031
QY 793 GAAGCGTGTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGA 852
Db 3032 GAAGCGTGTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGA 3091
QY 853 ATTAAGCCGGCAGGCACCTATTGAGGGCGACTTTCCAGCAGTTTCAGAAGCAGAACGATGT 912
Db 3092 ATTAAGCCGGCAGGCACCTATTGAGGGCGACTTTCCAGCAGTTTCAGAAGCAGAACGATGT 3151
QY 913 ACATAGGGCTTCAAGAGGGAATTGAAAACCTGTAATCATGAGTACTCTTGA 972
Db 3152 ACATAGGGCTTCAAGAGGGAATTGAAAACCTGTAATCATGAGTACTCTTGA 3211
QY 973 GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACTCTACCCAGGA 1032
Db 3212 GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACTCTACCCAGGA 3271
QY 1033 GCCCAGAGAGTGCCTCCTGTAGGAGAGAGCCCGAGAAATGTCACTCGGCTTCTACGAAAGCA 1092
Db 3272 GCCCAGAGAGTGCCTCCTGTAGGAGAGAGCCCGAGAAATGTCACTCGGCTTCTACGAAAGCA 3331
QY 1093 GGCTGAGGAGTCAATACTAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 1152
Db 3332 GGCTGAGGAGTCAATACTAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 3391
QY 1153 AAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGTGGGA 1212
Db 3392 AAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGTGGGA 3451
QY 1213 CCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCT 1272
Db 3452 CCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCT 3511
QY 1273 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTCGAGGAGAAATTGGGCC 1332
Db 3512 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTCGAGGAGAAATTGGGCC 3571
QY 1333 TCTGAAAGAGAACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAT 1392
Db 3572 TCTGAAAGAGAACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAT 3631
QY 1393 TCAGCTCTCACCGTATAACCTCAGCAGCTCTGGAAGACCTGAAACACCATGGAAGCTTCT 1452
Db 3632 TCAGCTCTCACCGTATAACCTCAGCAGCTCTGGAAGACCTGAAACACCATGGAAGCTTCT 3691
QY 1453 GCAGGTGGCCGCTGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGCACTTTGGTCC 1512

Db 3692 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGTGCATGAAGCCACAGGGACTTTGGTCC 3751
QY 1513 AGCATCTCAGCACATTTCTTCCACGTCTGTCCAGGTCCTGGGAGAGAGCCATCTCGCC 1572
Db 3752 AGCATCTCAGCACATTTCTTCCACGTCTGTCCAGGTCCTGGGAGAGAGCCATCTCGCC 3811
QY 1573 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACAACTTGTGGGACCATCCCAA 1632
Db 3812 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACAACTTGTGGGACCATCCCAA 3871
QY 1633 AATGACAGAGCTCTACCCAGTCTTTAGCTGACCTGAATAATGTGAGATTCTCAGCTTATAG 1692
Db 3872 AATGACAGAGCTCTACCCAGTCTTTAGCTGACCTGAATAATGTGAGATTCTCAGCTTATAG 3931
QY 1693 GACTGCCATGAAACTCCGAGAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 1752
Db 3932 GACTGCCATGAAACTCCGAGAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 3991
QY 1753 AGCTGCCATGTGATGCCCTTGACCCAGCACAACTTATGACCCGCTGGAGCAAGACACAA 1812
Db 3992 AGCTGCCATGTGATGCCCTTGACCCAGCACAACTTATGACCCGCTGGAGCAAGACACAA 4051
QY 1813 CCTGCAGATTATTAAATTTGTTGACCACTATTTATGACCCGCTGGAGCAAGACACAA 1872
Db 4052 CCTGCAGATTATTAAATTTGTTGACCACTATTTATGACCCGCTGGAGCAAGACACAA 4111
QY 1873 TTTGTCACACGTCCCTCTCTGCGTGGATATGTGTCGAACTGGCTGTGATGTTTATGA 1932
Db 4112 TTTGTCACACGTCCCTCTCTGCGTGGATATGTGTCGAACTGGCTGTGATGTTTATGA 4171
QY 1933 TACGGACGAAACAGGAGGATCCGTCCTGTCTTTTAAAACTGGCATCATTTCCCTGTG 1992
Db 4172 TACGGACGAAACAGGAGGATCCGTCCTGTCTTTTAAAACTGGCATCATTTCCCTGTG 4231
QY 1993 TAAAGCACATTTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 2052
Db 4232 TAAAGCACATTTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 4291
QY 2053 ATTTTGTGACCGCAGCGAGGCTGGGCTCCTTCTGATGATCTATCCAA 2101
Db 4292 ATTTTGTGACCGCAGCGAGGCTGGGCTCCTTCTGATGATCTATCCAA 4340

RESULT 10
US-09-845-416-32
; Sequence 32, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 32
; LENGTH: 4414
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-32

Query Match 78.1%; Score 1641; DB 13; Length 4414;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1848; Conservative 0; Mismatches 190; Indels 63; Gaps 3;
QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCCTGTGACCCCTACACGGAGCCCAT 60
Db 1657 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCCTGTGACCCCTACACGGAGCCCAT 1716
QY 61 TCCTTCACAGCATTTTGAAGCTCCTGAAAGCAAGTCAATTGGCAGTTTATTGATGGAGAG 120

Db 1717 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCAATTGATGGAGAG 1776
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
Db 1777 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1836
QY 181 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1837 TGCTGAGGACACATTTGCAAGCACAAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1896
QY 241 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 300
Db 1897 CCAGTTTTCATACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGTTGG 1956
QY 301 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1957 TAATATTCTACAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAGAAGATGAAGA 2016
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCGAAGATGGGAATGCCTCAGGGTAGC 420
Db 2017 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCGAAGATGGGAATGCCTCAGGGTAGC 2076
QY 421 TAGCATGGAaaaaCAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 480
Db 2077 TAGCATGGAaaaaCAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAACT 2136
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAAACAAGGAAAAATGGAGGAAGA 540
Db 2137 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAAACAAGGAAAAATGGAGGAAGA 2196
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAAACAACATAAGGTGCTTCA 600
Db 2197 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAAACAACATAAGGTGCTTCA 2256
QY 601 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGGTAGT 660
Db 2257 AGAAGATCTAGAAACAAGAACAAAGTCAGGGTCAATTTCTCACTCACATGGTGGGTAGT 2316
QY 661 TGATGAATCTPAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
Db 2317 TGATGAATCTPAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 2376
QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGTGGGTTCTTTTACAGACAG 780
Db 2377 AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGTGGGTTCTTTTACAGACAT 2436
QY 781 TTCTGACCACTGGAAGCGTCTGCACCTTTCTCTGCAGGAAGTCTTGGTGTGGCTACAGCT 840
Db 2437 CCTTCTCAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 2496
QY 841 GAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCGGACTTTCCAGCAGTTCAGAA 900
Db 2497 AAAAGAGATGCAGTGAACAAGATTACACAACTGGCTTTAAAGATCAAAATGAAATGTT 2556
QY 901 GCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCAT 960
Db 2557 ATCAAGTCTTCAAAAACTGGCCGTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 2616
QY 961 GAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAA 1020
Db 2617 GGGCAAACTGTA-----TTCACTCAAAACAAGATCTTCTTCAACACTGAAG 2662
QY 1021 ACTCTACAGGAGCCACAGAGAGTGCCTCCTGAGGAGAGAGCCCAGAAATGTCACTCGGCT 1080
Db 2663 AATAAGTCAGTGACCCAGAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGTGGG-- 2720
QY 1081 TCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTGCACCTCCGC 1140
Db 2721 -----ATAATTAGTCCAAAAAAGTTGAA----- 2743
QY 1141 TGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCAC 1200

Db 2744 -----AAGAGTACAGCACAGACCCCTTTGAAAAGACTCCAGGAACTTCAAGAGGCCAC 2793
QY 1201 GGATGAGCTGGACCTCAAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT 1260
Db 2794 GGATGAGCTGGACCTCAAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT 2853
QY 1261 GGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGG 1320
Db 2854 GGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGG 2913
QY 1321 AGAAATTGCGCCTCTGAAAGAGAAAGTGAAGCCACGTCATATGACCTTGTCTGCCAGCTTAC 1380
Db 2914 AGAAATTGCGCCTCTGAAAGAGAAAGTGAAGCCACGTCAGTCAATGACCTTGTCTGCCAGCTTAC 2973
QY 1381 CACTTTGGGCATTCAGCTCTCACCGTATAAACCCTCAGCACTCTGGAAGACCTGAACACCAG 1440
Db 2974 CACTTTGGGCATTCAGCTCTCACCGTATAAACCCTCAGCACTCTGGAAGACCTGAACACCAG 3033
QY 1441 ATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAG 1500
Db 3034 ATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAG 3093
QY 1501 GGACTTTTGGTCCAGCATCTCAGCACTTCTTTCCACGCTGTCTCCAGGGTCCCTGGGAGAG 1560
Db 3094 GGACTTTTGGTCCAGCATCTCAGCACTTCTTTCCACGCTGTCTCCAGGGTCCCTGGGAGAG 3153
QY 1561 AGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCACGAGACTCAAAACAACCTTGCTG 1620
Db 3154 AGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCACGAGACTCAAAACAACCTTGCTG 3213
QY 1621 GGACCATCCCAAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATT 1680
Db 3214 GGACCATCCCAAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATT 3273
QY 1681 CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGAGAAAGGCCCTTTTGTCTTGGATCT 1740
Db 3274 CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGAGAAAGGCCCTTTTGTCTTGGATCT 3333
QY 1741 CTTGAGCCTGTGAGTGCATGTGATGCCCTTGGACCAGCAACCTCAAGCAAAAATGACCA 1800
Db 3334 CTTGAGCCTGTGAGTGCATGTGATGCCCTTGGACCAGCAACCTCAAGCAAAAATGACCA 3393
QY 1801 GCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTATGACCGCCTGGAGCA 1860
Db 3394 GCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTATGACCGCCTGGAGCA 3453
QY 1861 AGAGCAAAACAATTTGGTCAACGTCCTCTCTCGTGGATATGTGTCTGAACTGGCTGCT 1920
Db 3454 AGAGCAAAACAATTTGGTCAACGTCCTCTCTCGTGGATATGTGTCTGAACTGGCTGCT 3513
QY 1921 GAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTCCTCTGCTTTTAAACTGGCAT 1980
Db 3514 GAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTCCTCTGCTTTTAAACTGGCAT 3573
QY 1981 CATTTCCTGTGTAAAGCACATTTTGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGC 2040
Db 3574 CATTTCCTGTGTAAAGCACATTTTGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGC 3633
QY 2041 AAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCTCCTTCTGTCATGATTCTATCCA 2100
Db 3634 AAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCTCCTTCTGTCATGATTCTATCCA 3693
QY 2101 A 2101
Db 3694 A 3694

RESULT 11
US-09-845-416-14
; Sequence 14, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO

; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: DE1142

; CURRENT APPLICATION NUMBER: US/09/845,416

; CURRENT FILING DATE: 2001-04-30

; PRIOR APPLICATION NUMBER: 60/200,777

; PRIOR FILING DATE: 2000-04-28

; NUMBER OF SEQ ID NOS: 36

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 14

; LENGTH: 3446

; TYPE: DNA

; ORGANISM: Homo sapiens

us-09-845-416-14

Query Match 77.5%; Score 1629; DB 13; Length 3446;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 1847; Conservative 0; Mismatches 190; Indels 64; Gaps 4;

| | | | |
|----|------|--|------|
| Qy | 1 | GAGCTATGCCCTACACACAGGCTGCTTATGTCACCCACCTCTGACCCCTACACGAGGCCCAT | 60 |
| Db | 900 | GAGCTATGCCCTACACACAGGCTGCTTATGTCACCCACCTCTGACCCCTACACGAGGCCCAT | 959 |
| Qy | 61 | TCCTTCACAGCATTTGGAAGCTCTCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG | 120 |
| Db | 960 | TCCTTCACAGCATTTGGAAGCTCTCTGAAGACAAGTCATTTGGCAGTTTCATTGATGGAGAG | 1019 |
| Qy | 121 | TGAACTAAACCTGGACCGTTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC | 180 |
| Db | 1020 | TGAACTAAACCTGGACCGTTTATCAAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC | 1079 |
| Qy | 181 | TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTTCTAATCATGTGTGGAAGTGGTGAAGA | 240 |
| Db | 1080 | TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTTCTAATCATGTGTGGAAGTGGTGAAGA | 1139 |
| Qy | 241 | CCAGTTTCATACACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGSCGGGTTGG | 300 |
| Db | 1140 | CCAGTTTCATACACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGSCGGGTTGG | 1199 |
| Qy | 301 | TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA | 360 |
| Db | 1200 | TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA | 1259 |
| Qy | 361 | AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCCTCAGGGTAGC | 420 |
| Db | 1260 | AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCCTCAGGGTAGC | 1319 |
| Qy | 421 | TAGCATGGAAAAACAAGCAATTTACATAGATTTTAAATGGATCTCCAGAAATCAGAAACT | 480 |
| Db | 1320 | TAGCATGGAAAAACAAGCAATTTACATAGATTTTAAATGGATCTCCAGAAATC-GAAACT | 1378 |
| Qy | 481 | GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAGGAAAATGGAGGAAGA | 540 |
| Db | 1379 | GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAGGAAAATGGAGGAAGA | 1438 |
| Qy | 541 | GCCTCTTGGACCTGATCTTGAGACCTAAAAACGCAAGTACAAACAACATAAGGTGCTTCA | 600 |
| Db | 1439 | GCCTCTTGGACCTGATCTTGAGACCTAAAAACGCAAGTACAAACAACATAAGGTGCTTCA | 1498 |
| Qy | 601 | AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTCTCTCACTCACATGGTGGGTAGT | 660 |
| Db | 1499 | AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTCTCTCACTCACATGGTGGGTAGT | 1558 |
| Qy | 661 | TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAAACAACCTTAAGGTATTGGG | 720 |
| Db | 1559 | TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAAACAACCTTAAGGTATTGGG | 1618 |
| Qy | 721 | AGATCGATGGCAAAACATCTGTAGATGGACAGAGCCGCTGGGTTCTTTTACAAGACAG | 780 |
| Db | 1619 | AGATCGATGGCAAAACATCTGTAGATGGACAGAGCCGCTGGGTTCTTTTACAAGACAT | 1678 |
| Qy | 781 | TTCTGACCAGTGGAAAGCGTCTGCACCTTCTCTGCAGGAACCTTCTGGTGTGCTACAGCT | 840 |

| | | | |
|----|------|---|------|
| Db | 1679 | CCTTCTCAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTGTAGTCAATGGCTTTTCTAGA | 1738 |
| Qy | 841 | GAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCTAGAA | 900 |
| Db | 1739 | AAAAGAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTTAAAGATCAAAATGAAATGTT | 1798 |
| Qy | 901 | GCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAAATCAT | 960 |
| Db | 1799 | ATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT | 1858 |
| Qy | 961 | GAGTACTCTTGAGACTGTACGAATATTTCTGCACAGCAGCCTTTTGGAAAGGACTAGAGAA | 1020 |
| Db | 1859 | GGCAAAACCTGTA-----TTCACTCAAAACAAGATCTTCTTTCAACACTGAAG | 1904 |
| Qy | 1021 | ACTCTACCAGGAGCCCGAGAGAGCTGCCTCTCTGAGGAGAGAGCCAGAAATGTCACTCGGT | 1080 |
| Db | 1905 | AATAAGTCAGTGACCCGAGAAAGACGGAAGCATGCGCTGGA----- | 1942 |
| Qy | 1081 | TCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTGCACTCCGC | 1140 |
| Db | 1943 | -----TAACTTTGGCCCGTGTGGGATAATTTAGTCCAAAAAATTGAA----- | 1985 |
| Qy | 1141 | TCAGTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCCAC | 1200 |
| Db | 1986 | -----AAGAGTACAGCACAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCCAC | 2035 |
| Qy | 1201 | GGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT | 1260 |
| Db | 2036 | GGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT | 2095 |
| Qy | 1261 | GGCGATCTCCTCATTGACTCTCTCCAGATCACCTCGAGAAAAGTCAAGGCACTTCGAGG | 1320 |
| Db | 2096 | GGCGATCTCCTCATTGACTCTCTCCAGATCACCTCGAGAAAAGTCAAGGCACTTCGAGG | 2155 |
| Qy | 1321 | AGAAATTGCGCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTAC | 1380 |
| Db | 2156 | AGAAATTGCGCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTAC | 2215 |
| Qy | 1381 | CACCTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAG | 1440 |
| Db | 2216 | CACCTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAG | 2275 |
| Qy | 1441 | ATGGAAGCTTCTGAGGTGGCCCGTTCGAGGACCGAGTCAAGGCTGCTGATGAAGCCACAG | 1500 |
| Db | 2276 | ATGGAAGCTTCTGAGGTGGCCCGTTCGAGGACCGAGTCAAGGCTGCTGATGAAGCCACAG | 2335 |
| Qy | 1501 | GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTGTGTCCAGGTCCTGGGAGAG | 1560 |
| Db | 2336 | GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTGTGTCCAGGTCCTGGGAGAG | 2395 |
| Qy | 1561 | AGCCATCTGCCAAAACAAGTGCCCTACTATATCAACACCAGAGACTCAAAACACTTGCTG | 1620 |
| Db | 2396 | AGCCATCTGCCAAAACAAGTGCCCTACTATATCAACACCAGAGACTCAAAACACTTGCTG | 2455 |
| Qy | 1621 | GGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATATGTGAGATT | 1680 |
| Db | 2456 | GGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATATGTGAGATT | 2515 |
| Qy | 1681 | CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCT | 1740 |
| Db | 2516 | CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCT | 2575 |
| Qy | 1741 | CTTGAGCCCTGTGAGTGTGATGCCCTTGACCCAGCAACCTCAAGCAAAATGACCA | 1800 |
| Db | 2576 | CTTGAGCCCTGTGAGTGTGATGCCCTTGACCCAGCAACCTCAAGCAAAATGACCA | 2635 |
| Qy | 1801 | GCCCATGGATATCCTGCAGATTATTAATTGTTTGGACCACTATTATGACCCGCTGGAGCA | 1860 |
| Db | 2636 | GCCCATGGATATCCTGCAGATTATTAATTGTTTGGACCACTATTATGACCCGCTGGAGCA | 2695 |
| Qy | 1861 | AGAGCACAACAATTTGGTCAACGTCCTCTCTGCGTGGATATGTGTCTGAACCTGGCTGCT | 1920 |
| Db | 2696 | AGAGCACAACAATTTGGTCAACGTCCTCTCTGCGTGGATATGTGTCTGAACCTGGCTGCT | 2755 |

| | | | | |
|----|------|----------------------|--|------|
| QY | 1921 | GAATGTTTATGATA | CGGGACGAACAGGAGGATCCGTGTCTTCTTTTAAACTGGCAT | 1980 |
| | | | | |
| | | | | |
| Db | 2756 | GAATGTTTATGATA | CGGGACGAACAGGAGGATCCGTGTCTTCTTTTAAACTGGCAT | 2815 |
| | | | | |
| | | | | |
| QY | 1981 | CATTTCCCTGTGTA | AAGCACATTTGGAGACAAGTACAGATACCTTTTCAAGCAAGTGGC | 2040 |
| | | | | |
| | | | | |
| Db | 2816 | CATTTCCCTGTGTA | AAAGCACATTTGGAGACAAGTACAGATACCTTTTCAAGCAAGTGGC | 2875 |
| | | | | |
| | | | | |
| QY | 2041 | AAGTTC AACAGGATTTTGT | GACCAAGCGCAGGCTGGGCCCTCCTTCTGTCATGATTCATATCCA | 2100 |
| | | | | |
| | | | | |
| Db | 2876 | AAGTTC AACAGGATTTTGT | GACCAAGCGCAGGCTGGGCCCTCCTTCTGTCATGATTCATATCCA | 2935 |
| | | | | |
| | | | | |
| QY | 2101 | A | 2101 | |
| | | | | |
| Db | 2936 | A | 2936 | |

RESULT 12

US-10-149-736-41

; Sequence 41, Application US/10149736

Publication No. US20030216332A1

GENERAL INFORMATION:

APPLICANT: Chamberlain, Jeffrey S.

APPLICANT: Harper, Scott Q.

TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences

FILE REFERENCE: UM-06968

; CURRENT APPLICATION NUMBER: US/10/149,736

CURRENT FILING DATE: 2002-06-17

PRIOR APPLICATION NUMBER: PCT/US01/311126

PRIOR FILING DATE: 2001-10-04

; PRIOR APPLICATION NUMBER: 60/238,848

PRIOR FILING DATE: 2000-10-06

; NUMBER OF SEQ ID NOS: 96

```

; SOFTWARE: PatentIn version 3.1

```

SEQ ID NO 41

; LENGTH: 5462

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Synthetic

US-10-149-736-41

| | | | | |
|----------------------------|--------|-----------------|-------------|--------------|
| Query Match | 77.5%; | Score 1629; | DB 13; | Length 5462; |
| Best Local Similarity | 85.4%; | Pred. No. 0; | | |
| Matches 1911; Conservative | 0; | Mismatches 190; | Indels 138; | Gaps 3; |

| | | | |
|----|------|--|------|
| QY | 1 | GAGCTATGCCTACACACAGGCTGCTTATGTCTACACACCTCTGACCCCTACACGAGGCCATT | 60 |
| DB | 1099 | GAGCTATGCCTACACACAGGCTGCTTATGTCTACACACCTCTGACCCCTACACGAGGCCATT | 1158 |
| QY | 61 | TCCTTCACAGCATTTGGAAAGCTCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG | 120 |
| DB | 1159 | TCCTTCACAGCATTTGGAAAGCTCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG | 1218 |
| QY | 121 | TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAAGTATTATCGTGGCTCTTTC | 180 |
| DB | 1219 | TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAAGTATTATCGTGGCTCTTTC | 1278 |
| ZY | 181 | TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA | 240 |
| Zb | 1279 | TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA | 1338 |
| ZY | 241 | CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG | 300 |
| Db | 1339 | CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG | 1398 |
| ZY | 301 | TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA | 360 |
| Db | 1399 | TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA | 1458 |
| ZY | 361 | AACTGAAGTACAAAGACAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC | 420 |

QY 959 ATGAGTACTCTTGAGACTGTACGAATATTT-----CTGACAGAGCAGCCTTTTGA 1008
Db 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTGAC 2118
QY 1009 AGGACTAGAGAAACTCTACCAGGAGCCACAGAGCTGCCTCTCCTGAGGAGAGAGCCAGAA 1068
Db 2119 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCGGTGTGGGATAATTTAGTCCAAA 2178
QY 1069 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATAC----- 1110
Db 2179 ACTTGAAAAGAGTACAGCACAGATTTACAGGCTGTCCACCACCTCAGCCATCACTAAC 2238
QY 1111 ----- 1110
Db 2239 ACAGACAACCTGTAATGGAACACAGTAACACTACGGTGACCACAGGGAACAGATCCTGGTAAA 2298
QY 1111 ---TGAGTGGGAAAAATTGAACCTTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGAC 1167
Db 2299 GCATGCTCAAGAGGAACCTTCCACCACCACTCCCCAAAAGAGAGGCAGATTACTGTGGA 2358
QY 1168 CCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCA 1227
Db 2359 TCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCA 2418
QY 1228 AGCTGAGTGATCAAGGGATCTTGGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCCA 1287
Db 2419 AGCTGAGTGATCAAGGGATCTTGGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCCA 2478
QY 1288 AGATCACTCGAGAAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAAGAGAAAGT 1347
Db 2479 AGATCACTCGAGAAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAAGAGAAAGT 2538
QY 1348 GAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTA 1407
Db 2539 GAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTA 2598
QY 1408 TAACCTCAGCACTCTGGAAGACCTGAAACACCAGATGGAAGCTTCTGCAGGTGGCGTGA 1467
Db 2599 TAACCTCAGCACTCTGGAAGACCTGAAACACCAGATGGAAGCTTCTGCAGGTGGCGTGA 2658
QY 1468 GGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTT 1527
Db 2659 GGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTT 2718
QY 1528 TCTTTCCAGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAGTGCCTTA 1587
Db 2719 TCTTTCCAGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAGTGCCTTA 2778
QY 1588 CTATATCAACCCAGAGACTCAAAACAATTGCTGGGACCATCCCAAAATGACAGAGCTCTA 1647
Db 2779 CTATATCAACCCAGAGACTCAAAACAATTGCTGGGACCATCCCAAAATGACAGAGCTCTA 2838
QY 1648 CCAGCTTTTAGCTGACCTGAAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACT 1707
Db 2839 CCAGCTTTTAGCTGACCTGAAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACT 2898
QY 1708 CCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGATGTGATGC 1767
Db 2899 CCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGATGTGATGC 2958
QY 1768 CTTGGACCGACACAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAATTAATAA 1827
Db 2959 CTTGGACCGACACAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAATTAATAA 3018
QY 1828 TTGTTTGACCACTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCC 1887
Db 3019 TTGTTTGACCACTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCC 3078
QY 1888 TCTCTGCGTGGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGG 1947
Db 3079 TCTCTGCGTGGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGG 3138

QY 1948 GAGGATCCGTCTCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGA 2007
Db 3139 GAGGATCCGTCTCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGA 3198
QY 2008 AGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCG 2067
Db 3199 AGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCG 3258
QY 2068 CAGGCTGGGCTCTCTTCTGCAATGATTTCTATCCAA 2101
Db 3259 CAGGCTGGGCTCTCTTCTGCAATGATTTCTATCCAA 3292

RESULT 14

US-09-845-416-6

; Sequence 6, Application US/09845416

; Publication No. US20030171312A1

; GENERAL INFORMATION:

; APPLICANT: XIAO, XIAO

; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE

; FILE OF INVENTION: THEREOF

; FILE REFERENCE: DE1142

; CURRENT APPLICATION NUMBER: US/09/845,416

; CURRENT FILING DATE: 2001-04-30

; PRIOR APPLICATION NUMBER: 60/200,777

; PRIOR FILING DATE: 2000-04-28

; NUMBER OF SEQ ID NOS: 36

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 3999

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-845-416-6

Query Match 76.2%; Score 1602; DB 13; Length 3999;

Best Local Similarity 81.1%; Pred. No. 0;

Matches 2101; Conservative 0; Mismatches 0; Indels 489; Gaps 1;

QY 1 GAGCTATGCTTACACACAGGCTGCTTATGTACCACTCTGACCCCTACACGGAGCCCAT 60
Db 900 GAGCTATGCTTACACACAGGCTGCTTATGTACCACTCTGACCCCTACACGGAGCCCAT 959
QY 61 TCCTTACAGCATTTGGAAGCTCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 120
Db 960 TCCTTACAGCATTTGGAAGCTCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 1019
QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAATTTATCGTGGCTTCTTC 180
Db 1020 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAAATTTATCGTGGCTTCTTC 1079
QY 181 TGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1080 TGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1139
QY 241 CCAGTTTCATCTCATGAGGCTACATGATGGATTTGACAGCCCATCAGGCGCGGTTGG 300
Db 1140 CCAGTTTCATCTCATGAGGCTACATGATGGATTTGACAGCCCATCAGGCGCGGTTGG 1199
QY 301 TAATATTCTACAATTGGGAGTAAGCTGATTGGAAACAGGAAAATTTATCAGAAGATGAAGA 360
Db 1200 TAATATTCTACAATTGGGAGTAAGCTGATTGGAAACAGGAAAATTTATCAGAAGATGAAGA 1259
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGTTAGC 420
Db 1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGTTAGC 1319
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGATTTTAAATGGATCTCCAGATCAGAAACT 480
Db 1320 TAGCATGGAAAAACAAAGCAATTTACATAGATTTTAAATGGATCTCCAGATCAGAAACT 1379
QY 481 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAAGGAAAATGGAGGAAGA 540
Db 1380 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAAGGAAAATGGAGGAAGA 1439

QY 541 GCCTCTTGACCTGATCTTGAAGACCTAAACGCCAAGTACAACAACATAAAGTGCTTCA 600
Db 1440 GCCTCTTGACCTGATCTTGAAGACCTAAACGCCAAGTACAACAACATAAAGTGCTTCA 1499
QY 601 AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT 660
Db 1500 AGAAGATCTAGAACAAAGCAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT 1559
QY 661 TGATGAATCTAGTGGAGATCAAGCAACTGCTGCTTTGGAGAAACAACCTTAAGGTATGGG 720
Db 1560 TGATGAATCTAGTGGAGATCAAGCAACTGCTGCTTTGGAGAAACAACCTTAAGGTATGGG 1619
QY 721 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGA--- 777
Db 1620 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACCA 1679
QY 778 ----- 777
Db 1680 GCCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTCTCTACTCAGACTGTTACTCT 1739
QY 778 ----- 777
Db 1740 GGTGACACAACCTGTGGTTACTAAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTTC 1799
QY 778 ----- 777
Db 1800 CTTGATGTTGAGGTACCTACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAA 1859
QY 778 ----- 777
Db 1860 GTTTCCTTGCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCG 1919
QY 778 ----- 777
Db 1920 TAAGGAAGGCTCCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGSCAAGA 1979
QY 778 ----- 777
Db 1980 CCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAGCCCA 2039
QY 778 ----- 777
Db 2040 AAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTACAAAGACGTTTGA 2099
QY 778 ----- 777
Db 2100 TAACATGAATCTCAAGTGGAGTGAATTCGGAAAAAGTCTCTCAACATTAGTCCCATTT 2159
QY 778 ----- CAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTG 831
Db 2160 GGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTG 2219
QY 832 GCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTTGGAGGCGACTTTCCAGC 891
Db 2220 GCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTTGGAGGCGACTTTCCAGC 2279
QY 892 AGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAAC 951
Db 2280 AGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAAC 2339
QY 952 TGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGAAGG 1011
Db 2340 TGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGAAGG 2399
QY 1012 ACTAGAGAAACTCTPACAGGAGCCCGAGAGCTGCCTCTCTGAGGAGAGAGCCCGAGATGT 1071
Db 2400 ACTAGAGAAACTCTPACAGGAGCCCGAGAGCTGCCTCTCTGAGGAGAGAGCCCGAGATGT 2459
QY 1072 CACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACTT 1131
Db 2460 CACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACTT 2519

QY 1132 GCATCCGCTGACTGGCAGAGAAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCA 1191
Db 2520 GCATCCGCTGACTGGCAGAGAAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCA 2579
QY 1192 AGAGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTG 1251
Db 2580 AGAGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTG 2639
QY 1252 GCAGCCGCTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGC 1311
Db 2640 GCAGCCGCTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGC 2699
QY 1312 ACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTCAGCCACGTCATGACCTTGCTCG 1371
Db 2700 ACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTCAGCCACGTCATGACCTTGCTCG 2759
QY 1372 CCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCCGTATAACCTCAGCACTCTGGAGACCT 1431
Db 2760 CCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCCGTATAACCTCAGCACTCTGGAGACCT 2819
QY 1432 GAACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGACGTCGATGA 1491
Db 2820 GAACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGACGTCGATGA 2879
QY 1492 AGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTTCCACGCTCTGTCCAGGGTCC 1551
Db 2880 AGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTTCCACGCTCTGTCCAGGGTCC 2939
QY 1552 CTGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCAGAGACTCAAAC 1611
Db 2940 CTGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCAGAGACTCAAAC 2999
QY 1612 AACTTGCTGGACCATCCCAAAATGACAGAGCTCTACCAGCTTTTAGCTGACCTGAATAA 1671
Db 3000 AACTTGCTGGACCATCCCAAAATGACAGAGCTCTACCAGCTTTTAGCTGACCTGAATAA 3059
QY 1672 TGTCAGATTCTCAGCTTATAGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTG 1731
Db 3060 TGTCAGATTCTCAGCTTATAGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTG 3119
QY 1732 CTTGGATCTCTTGAGCCTGTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 1791
Db 3120 CTTGGATCTCTTGAGCCTGTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 3179
QY 1792 AAATGACAGCCCATGGATATCCTGCAGATTATTAAATGTTTGGACCACTATTATGACCG 1851
Db 3180 AAATGACAGCCCATGGATATCCTGCAGATTATTAAATGTTTGGACCACTATTATGACCG 3239
QY 1852 CCTGGAGCAAGAGCACAAACAAATTTGGTCAACGTCCTCTCTGCGTGGATATGTCGTGAA 1911
Db 3240 CCTGGAGCAAGAGCACAAACAAATTTGGTCAACGTCCTCTCTGCGTGGATATGTCGTGAA 3299
QY 1912 CTGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGCTTTTAA 1971
Db 3300 CTGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGCTTTTAA 3359
QY 1972 AACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAA 2031
Db 3360 AACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAA 3419
QY 2032 GCAAGTGCAGATTCAACAGGATTTGTGACCAGCGCAGGCTGGGCTCCTTCTGTCATGA 2091
Db 3420 GCAAGTGCAGATTCAACAGGATTTGTGACCAGCGCAGGCTGGGCTCCTTCTGTCATGA 3479
QY 2092 TTCTATCCAA 2101
Db 3480 TTCTATCCAA 3489

; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 4966
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-28

Query Match 76.2%; Score 1602; DB 13; Length 4966;
Best Local Similarity 81.1%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 489; Gaps 1;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTACCCACCTCTGACCCCTACACGGAGCCCAT 60
Db |||||||
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG 120
Db |||||||
QY 1717 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAAGTCATTTGGCAGTTCATTGATGGAGAG 1776
Db |||||||
QY 121 TGAAGTAACTGGACCGTTATCAACACAGCTTTAGAAAGATATTATCGTGGCTTCTTTC 180
Db |||||||
QY 1777 TGAAGTAACTGGACCGTTATCAACACAGCTTTAGAAAGATATTATCGTGGCTTCTTTC 1836
Db |||||||
QY 181 TGCTGAGGACACATTGCAAGCAACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db |||||||
QY 1837 TGCTGAGGACACATTGCAAGCAACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1896
Db |||||||
QY 241 CCAGTTTCAFACTCATGAGGGGTACATGATGGATTGACAGGCCATCAGGGCCGGTTGG 300
Db |||||||
QY 1897 CCAGTTTCAFACTCATGAGGGGTACATGATGGATTGACAGGCCATCAGGGCCGGTTGG 1956
Db |||||||
QY 301 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 360
Db |||||||
QY 1957 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 2016
Db |||||||
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCCTCAGGTTAGC 420
Db |||||||
QY 2017 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCCTCAGGTTAGC 2076
Db |||||||
QY 421 TAGCATGGAAGAAACAAAGCAATTTTACATAGAGTTTAAATGGATCTCCAGAATCAGAACT 480
Db |||||||
QY 2077 TAGCATGGAAGAAACAAAGCAATTTTACATAGAGTTTAAATGGATCTCCAGAATCAGAACT 2136
Db |||||||
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGAAATGGAGGAAGA 540
Db |||||||
QY 2137 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGAAATGGAGGAAGA 2196
Db |||||||
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCAAAGTACAAACATAAGGTGCTTCA 600
Db |||||||
QY 2197 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCAAAGTACAAACATAAGGTGCTTCA 2256
Db |||||||
QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAAGGGTCAATTTCTCACTCACATGGTGGTGTAGT 660
Db |||||||
QY 2257 AGAAGATCTAGAACAAAGAACAAAGTCAAGGGTCAATTTCTCACTCACATGGTGGTGTAGT 2316
Db |||||||
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAGAACAACTTAAGGTATTGGG 720
Db |||||||
QY 2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAGAACAACTTAAGGTATTGGG 2376
Db |||||||
QY 721 AGATCGATGGCAAAACATCTGTAGATGGACAGAGACCGTGGGTTCTTTTACAAGA--- 777
Db |||||||
QY 2377 AGATCGATGGCAAAACATCTGTAGATGGACAGAGACCGTGGGTTCTTTTACAAGACCA 2436
Db |||||||

QY 778 ----- 777
Db 2437 GCCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTCTCCTACTCAGACTGTTACTCT 2496
QY 778 ----- 777
Db 2497 GGTGACACAACTGTGGTTACTAAGGAAACTGCCATCTCCAAACTAGAAAATGCCATCTTC 2556
QY 778 ----- 777
Db 2557 CTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAGTTCGCCCTGCCACTTGAAAA 2616
QY 778 ----- 777
Db 2617 GTTTCTTGCTGGCTTACAGAAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCG 2676
QY 778 ----- 777
Db 2677 TAAGGAAAGGCTCCTAGAGACTCCRAAGGGAGTAAAGAGCTGATGAAACAAATGGCAAGA 2736
QY 778 ----- 777
Db 2737 CCTCCAAGGTGAAATGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAAAACAGCCA 2796
QY 778 ----- 777
Db 2797 AAAAATCCTGAGATCCCTTGGAAAGTTCGGATGATGCAGTCTCTGTTACAAAGACGTTTGG 2856
QY 778 ----- 777
Db 2857 TAACATGAACCTCAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAGGTCCCATTT 2916
QY 778 ----- 831
Db 2917 GGAAGCCAGTCTGACCAAGTGGAAAGCTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTG 2976
QY 832 GCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTTGGAGGCGACTTTCCAGC 891
Db |||||||
QY 2977 GCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTTGGAGGCGACTTTCCAGC 3036
QY 892 AGTTCAGAAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTTGAAAACTAAAGAAC 951
Db |||||||
QY 3037 AGTTCAGAAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTTGAAAACTAAAGAAC 3096
QY 952 TGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGGAAAG 1011
Db |||||||
QY 3097 TGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGGAAAG 3156
QY 1012 ACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCCTCTGAGGAGAGAGCCCGAGAAATGT 1071
Db |||||||
QY 3157 ACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCCTCTGAGGAGAGAGCCCGAGAAATGT 3216
QY 1072 CACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGAACCT 1131
Db |||||||
QY 3217 CACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGAACCT 3276
QY 1132 GCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAAGACTCCAGGAACCTCA 1191
Db |||||||
QY 3277 GCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTCA 3336
QY 1192 AGAGGCCACGGATGAGCTGGACCTCAAGCTCGGCCAAGCTGAGGTGATCAAGGGATCCTG 1251
Db |||||||
QY 3337 AGAGGCCACGGATGAGCTGGACCTCAAGCTCGGCCAAGCTGAGGTGATCAAGGGATCCTG 3396
QY 1252 GCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGAGAAAAGTCAAGGC 1311
Db |||||||
QY 3397 GCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGAGAAAAGTCAAGGC 3456
QY 1312 ACTTCGAGGAGAAAAATTCGCCCTCTGAAAGAGAACGCTGAGCCACGTCATGACCTTGTCTCG 1371
Db |||||||
QY 3457 ACTTCGAGGAGAAAAATTCGCCCTCTGAAAGAGAACGCTGAGCCACGTCATGACCTTGTCTCG 3516
QY 1372 CCAGCTTACCACCTTTGGGCAATTCAGCTCTCAACCGTATTAACCTCAGCACTCTGGAAGACCT 1431

Db 3517 CCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAAACCCTCAGCACTCTTGGAGACCT 3576
QY 1432 GAACACCATGATGGAAGCTTCTGACGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGA 1491
Db 3577 GAACACCATGATGGAAGCTTCTGACGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGA 3636
QY 1492 AGCCACACAGGACTTTGGTCCAGATCTCTCAGCACTTTCTTCCACGTCGTCTCCAGGGTCC 1551
Db 3637 AGCCACACAGGACTTTGGTCCAGATCTCTCAGCACTTTCTTCCACGTCGTCTCCAGGGTCC 3696
QY 1552 CTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCCACGAGACTCAAAC 1611
Db 3697 CTGGGAGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCCACGAGACTCAAAC 3756
QY 1612 AACTTGCTGGACCATCCCAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAA 1671
Db 3757 AACTTGCTGGACCATCCCAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAA 3816
QY 1672 TGTCAGATTCTCAGCTTATAGGACTGCGCATGAAACTCCGAAGACTGCAGAAGGCCCTTG 1731
Db 3817 TGTCAGATTCTCAGCTTATAGGACTGCGCATGAAACTCCGAAGACTGCAGAAGGCCCTTG 3876
QY 1732 CTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGACCAGCACAAACCTCAAGCA 1791
Db 3877 CTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGACCAGCACAAACCTCAAGCA 3936
QY 1792 AAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTTGTTGACCACTATTTATGACCG 1851
Db 3937 AAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTTGTTGACCACTATTTATGACCG 3996
QY 1852 CCTGGAGCAAGAGCACAACAATTTGTFCAACGTCCCTCTCTGGTGGATATGTGTCTGAA 1911
Db 3997 CCTGGAGCAAGAGCACAACAATTTGTFCAACGTCCCTCTCTGGTGGATATGTGTCTGAA 4056
QY 1912 CTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCGTCTTTAA 1971
Db 4057 CTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCGTCTTTAA 4116
QY 1972 AACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAA 2031
Db 4117 AACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAA 4176
QY 2032 GCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGCGCTGGGCCCTCCTTCTGCATGA 2091
Db 4177 GCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGCGCTGGGCCCTCCTTCTGCATGA 4236
QY 2092 TTCTATCCAA 2101
Db 4237 TTCTATCCAA 4246

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 11:04:35 ; Search time 3939.6 Seconds
(without alignments)
12961.661 Million cell updates/sec

Title: US-09-845-416-12_COPY_900_3000

Perfect score: 2101

Sequence: 1 gagctatgcctacacacagg.....ttctgcatgattctatccaa 2101

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estmu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_htc: *
9: gb_est1: *
10: gb_est2: *
11: gb_est3: *
12: gb_est4: *
13: gb_est5: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: em_gss_hum: *
18: em_gss_inv: *
19: em_gss_pln: *
20: em_gss_vrt: *
21: em_gss_fun: *
22: em_gss_mam: *
23: em_gss_mus: *
24: em_gss_pro: *
25: em_gss_rod: *
26: em_gss_phg: *
27: em_gss_vrl: *
28: gb_gss1: *
29: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 787 | 37.5 | 3870 | 11 BC036103 | BC036103 Homo sapi |
| 2 | 629.8 | 30.0 | 3056 | 11 AK044536 | AK044536 Mus muscu |
| 3 | 514.6 | 24.5 | 1490 | 11 BC009242 | BC009242 Homo sapi |
| 4 | 509.2 | 24.2 | 3753 | 11 AK081426 | AK081426 Mus muscu |

| | | | | | | | |
|---|----|-------|------|------|----|-----------|--------------------|
| C | 5 | 509 | 24.2 | 595 | 14 | CB177816 | CB177816 is21c01.x |
| | 6 | 505 | 24.0 | 824 | 9 | AL556247 | AL556247 AL556247 |
| | 7 | 503.4 | 24.0 | 801 | 14 | CB991394 | CB991394 AGENCOURT |
| | 8 | 503.4 | 24.0 | 1098 | 13 | BX365572 | BX365572 BX365572 |
| | 9 | 494 | 23.5 | 620 | 13 | BQ640063 | BQ640063 he23g04.y |
| | 10 | 494 | 23.5 | 797 | 14 | CB960722 | CB960722 AGENCOURT |
| | 11 | 480 | 22.8 | 728 | 14 | CB228986 | CB228986 AGENCOURT |
| | 12 | 471.4 | 22.4 | 770 | 10 | BG719710 | BG719710 602690430 |
| | 13 | 455.8 | 21.7 | 591 | 2 | HSM083521 | Bx490860 Homo sapi |
| | 14 | 453.8 | 21.6 | 652 | 10 | BB629984 | BB629984 BB629984 |
| | 15 | 453.8 | 21.6 | 1298 | 11 | AK087829 | AK087829 Mus muscu |
| | 16 | 453.8 | 21.6 | 2135 | 11 | AK013510 | AK013510 Mus muscu |
| | 17 | 453.8 | 21.6 | 4437 | 11 | AK036936 | AK036936 Mus muscu |
| | 18 | 452.2 | 21.5 | 1384 | 11 | AK075809 | AK075809 Mus muscu |
| | 19 | 449 | 21.4 | 704 | 10 | BB610411 | BB610411 BB610411 |
| | 20 | 439 | 20.9 | 578 | 2 | HSM075761 | Bx485574 Homo sapi |
| | 21 | 428.6 | 20.4 | 854 | 9 | AI196693 | AI196693 ui53e10.y |
| | 22 | 423 | 20.1 | 665 | 14 | BY742604 | BY742604 BY742604 |
| | 23 | 410.4 | 19.5 | 599 | 10 | BB666688 | BB666688 BB666688 |
| | 24 | 407.2 | 19.4 | 554 | 14 | CB613696 | CB613696 AMGNNUC:N |
| | 25 | 404.6 | 19.3 | 3051 | 11 | BC036095 | BC036095 Homo sapi |
| | 26 | 404.4 | 19.2 | 493 | 14 | CA888041 | CA888041 B0142C06- |
| | 27 | 398.2 | 19.0 | 717 | 14 | CB527785 | CB527785 UI-M-FY0- |
| | 28 | 397.4 | 18.9 | 495 | 14 | CA894775 | CA894775 B0187G06- |
| | 29 | 392.2 | 18.7 | 650 | 14 | BY714491 | BY714491 BY714491 |
| | 30 | 388.6 | 18.5 | 644 | 13 | BU313510 | BU313510 603540290 |
| | 31 | 383.4 | 18.2 | 483 | 14 | CA893902 | CA893902 B0182B01- |
| | 32 | 381.8 | 18.2 | 423 | 9 | AA460476 | AA460476 zx61e10.r |
| | 33 | 380.2 | 18.1 | 684 | 9 | AL641565 | AL641565 AL641565 |
| | 34 | 375.4 | 17.9 | 514 | 9 | AL871560 | AL871560 AL871560 |
| | 35 | 374.8 | 17.8 | 898 | 10 | BF182065 | BF182065 601804604 |
| | 36 | 369.6 | 17.6 | 843 | 14 | CA988247 | CA988247 AGENCOURT |
| | 37 | 368.6 | 17.5 | 633 | 9 | AL796733 | AL796733 AL796733 |
| | 38 | 349.6 | 16.6 | 546 | 4 | BX516597 | Bx516597 RZPD Mus |
| | 39 | 334.6 | 15.9 | 794 | 13 | BU424348 | BU424348 603234189 |
| | 40 | 332.6 | 15.8 | 541 | 9 | AL894729 | AL894729 AL894729 |
| | 41 | 330 | 15.7 | 608 | 9 | AL672616 | AL672616 AL672616 |
| | 42 | 329 | 15.7 | 532 | 10 | BE334408 | BE334408 ug97g06.y |
| | 43 | 326.2 | 15.5 | 655 | 14 | CA558919 | CA558919 K0251G03- |
| | 44 | 325.6 | 15.5 | 532 | 12 | BJ036615 | BJ036615 BJ036615 |
| C | 45 | 324.6 | 15.4 | 402 | 14 | CB547284 | CB547284 AMGNNUC:S |

ALIGNMENTS

| | | | | | |
|------------|--|--|-------------|--------|-----------------|
| RESULT 1 | BC036103 | 3870 bp | mRNA | linear | HTC 04-MAR-2003 |
| LOCUS | BC036103 | Homo sapiens, clone IMAGE:5274415, | mRNA. | | |
| DEFINITION | BC036103 | BC036103.1 | GI:23271310 | | |
| ACCESSION | BC036103 | HTC. | | | |
| VERSION | BC036103.1 | Homo sapiens (human) | | | |
| KEYWORDS | HTC. | Homo sapiens | | | |
| SOURCE | Homo sapiens | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; | | | |
| ORGANISM | Homo sapiens | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | |
| REFERENCE | 1 | (bases 1 to 3870) | | | |
| AUTHORS | Strausberg, R. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (31-JUL-2002) National Institutes of Health, Mammalian | | | | |
| | Gene Collection (MGC), Cancer Genomics Office, National Cancer | | | | |
| | Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, | | | | |
| | USA | | | | |
| REMARK | NIH-MGC Project URL: http://mgc.nci.nih.gov | | | | |
| COMMENT | Contact: MGC help desk | | | | |
| | Email: cgapbs@mail.nih.gov | | | | |
| | Tissue Procurement: Miklos Palkovits, M.D., Ph.D. | | | | |
| | CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki | | | | |
| | Toshiyuki and Piero Carninci (RIKEN) | | | | |
| | CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) | | | | |
| | DNA Sequencing by: Sequencing Group at the Stanford Human Genome | | | | |
| | Center, Stanford University School of Medicine, Stanford, CA 94305 | | | | |

Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
and Hayashizaki,Y.
Functional annotation of a full-length mouse cDNA collection

Nature 409 (6821), 685-690 (2001)
21085660
11217851
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)
6 (bases 1 to 3056)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Retina RNA was provided by Dr. Stefano Gustincich (Department of
Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA
02115, USA) whose assistance is gratefully acknowledged. Please
visit our web site for further details.

URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.

Location/Qualifiers
1..3056
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:A930019F21"
/db_xref="taxon:10090"
/clone="A930019F21"
/tissue type="retina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"

misc_feature
1..3056
/note="dystrophin, muscular dystrophy (MGD|MGI:94909,
GB|NM_007868, evidence: BLASTN, 100%, match=999)"

BASE COUNT 1072 a 567 c 607 g 810 t
ORIGIN

Query Match 30.0%; Score 629.8; DB 11; Length 3056;
Best Local Similarity 83.4%; Pred. No. 3.8e-121;
Matches 715; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 1 GAGCTATGCCTACACACAGGCTGCTTATGTCAACCACTCTGACCCCTACACGGAGCCCAATT 60
|||||

Db 944 GAGTTATGCCTTACACACAGGCTGCTTATGTGCGCACTCTGATTCACACAGAGCCCTA 1003
|||||

QY 61 TCCTTCACAGCATTGGGAAGCTCCTGAACACAAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
|||||

Db 1004 TCCTTCACAGCATTTTGGAAGCTCCAGAGACAAAGTCACTTGACAGTTTCATTGATGGAGAC 1063

QY 121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTCTTTTC 180
|||||

Db 1064 GGAAGTAAATCTGGATAGTTACCAAACCTGCTTTAGAAGAAGTACTTTTCATGGCTCTTTTC 1123
|||||

QY 181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTTAATGATGTGGAAGTGTGAAAGA 240
|||||

Db 1124 TGCCGAGGATACATTGCGAGCACAAGGAGAGATTTCAAATGATGTTGAAGAAGTGAAGA 1183
|||||

QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTCAGAGCCCATCAGGGCCGGTTGG 300
|||||

Db 1184 ACAGTTTCATCTCATGAGGGATTTCATGATGGATTTCATCATCAAGGACTTCTTTGG 1243
|||||

QY 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGACAGGAAAAATTATCAGAAGATGAAGA 360
|||||

Db 1244 TAATGTTCTACAGTTAGGAAGTCAACTAGTTGGAAAAAGGAAAAATTATCAGAAGATGAAGA 1303
|||||

QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGCTAGC 420
|||||

Db 1304 AGCTGAAGTGCAAAGAACAAATGAATCTCCTAAATTCAAGATGGGAATGTCTCAGGGGTAGC 1363
|||||

QY 421 TAGCATGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAAACT 480
|||||

Db 1364 TAGCATGAAAAACAAAGCAATTTACACAAAGTTCTTAATGGATCTCCAGAAATCAGAAAAAT 1423
|||||

QY 481 GAAAGATTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAAGGAAAAATGGAGGAAGA 540
|||||

Db 1424 AAAAGAACTAGATGACTGTGTTAAACAAAAACTGAAGAGAGAACTAAGAAAAATGGAGGAAGA 1483
|||||

QY 541 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCCAAGTACAACAAACATAAAGTGTCTTCA 600
|||||

Db 1484 GCCCTTGGACCTGATCTTGAAGATCTAAATGCCCAAGTACAACAAACATAAAGTGTCTTCA 1543
|||||

QY 601 AGAAGATCTAGAACAAAGCAAGTCAAGGTCAATTCTCTCACTCACATGGTGTGTGTAGT 660
|||||

Db 1544 AGAAGATCTAGAACAGGAGCAGGTCAAGGTCAACTCGCTCACTCACATGGTAGTGGT 1603
|||||

QY 661 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
|||||

Db 1604 TGATGAATCCAGCGGTGATCATGCAACAGCTGCTTTGGAAGAACAACTTAAGGTACTGGG 1663
|||||

QY 721 AGATCGATGGGCAACATCTCTAGATGGACAGAACGCCGTGGTTCCTTTTACAGACAG 780
|||||

Db 1664 AGATCGATGGGCAATATCTGCAGATGGACTGAAGACCGCTGGATTGTTTTACAGATAT 1723
|||||

QY 781 TTCTGACCACTGGAAGCGTCTGCACCTTTCTCTGACGAACTTCTGGTGTGGCTACAGCT 840
|||||

Db 1724 TCTTCTAAAATGGCAGCATTTTACTGAAGAACAGTGCCTTTTAGTACATGGCTTTCAGA 1783
|||||

QY 841 GAAAGATGATGAATTA 857
|||||

Db 1784 AAAAGAAGATGCAATGA 1800
|||||

RESULT 3
BC009242 1490 bp mRNA linear HTC 04-MAR-2003
LOCUS Homo sapiens, dystrophin (muscular dystrophy, Duchenne and Becker
types), includes DXS142, DXS164, DXS206, DXS230, DXS239, DXS268,
DXS269, DXS270, DXS272, clone IMAGE:3029414, mRNA.
BC009242
BC009242.1 GI:14714379
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1490)
Strausberg,R.
Direct Submission
Submitted (06-JUN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaps-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ruben Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbiology.org>
contact: amadani@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 10 Row: j Column: 10
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5032284
This clone has the following problem: retained intron.

FEATURES

source
1..1490
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3029414"
/tissue_type="Muscle, rhabdomyosarcoma"
/clone_lib="NIH MGC 17"
/lab_host="DH10B-R"
/note="Vector: pOTB7" 358 t

BASE COUNT 505 a 299 c 328 g 358 t
ORIGIN

Query Match. 24.5%; Score 514.6; DB 11; Length 1490;
Best Local Similarity 99.2%; Pred. No. 4.2e-97;
Matches 517; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GAGCTATGCCTACACAGGCTGCTTATGTCACCACCTCTGACCTTACACGGAGCCCAT 60
Db |||||
QY 969 GAGCTATGCCTACACAGGCTGCTTATGTCACCACCTCTGACCTTACACGGAGCCCAT 1028
Db |||||
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 120
Db |||||
QY 1029 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1088
Db |||||
QY 121 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAACTATTATCGTGGCTTCTTTC 180
Db |||||
QY 1089 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAACTATTATCGTGGCTTCTTTC 1148
Db |||||
QY 181 TGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db |||||
QY 1149 TGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1208
Db |||||
QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG 300
Db |||||
QY 1209 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG 1268
Db |||||
QY 301 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
Db |||||
QY 1269 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1328
Db |||||
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
Db |||||
QY 1329 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1388
Db |||||
QY 421 TAGCATGGAAAAACAAAGCAATTTCATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 480
Db |||||
QY 1389 TAGCATGGAAAAACAAAGCAATTTCATAGAGTTTAAATGGATCTCCAGAATCAGAAACT 1448
Db |||||
QY 481 GAAAGAGTTGAATGACTGCTTAACAAAAACAGAAAGAA 521
Db |||||
QY 1449 GAAAGAGTTGAATGACTGCTTAACAAAAACAGAAAGAA 1489
Db |||||

RESULT 4
AK081426
LOCUS
DEFINITION

AK081426 3753 bp mRNA linear HTC 05-DEC-2002
Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone:C130016K19 product:dystrophin related protein 2, full insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AK081426 GI:26349154

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE

1 Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

REFERENCE

AUTHORS

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

11042159

REFERENCE

AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,

Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,

Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system-384-format

sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

REFERENCE

AUTHORS

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,

Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,

Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,

Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,

Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,

Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,

Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,

Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,

Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,

Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,

Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,

Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,

Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,

Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,

Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,

Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,

Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,

Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.

and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection

Nature 409 (6821), 685-690 (2001)

21085660

11217851

REFERENCE

AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research

Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 3753)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,

Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE
JOURNAL

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/

FEATURES

source

1. 3753
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:C130016K19"
/db_xref="taxon:10090"
/clone="C130016K19"
/tissue_type="head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days embryo"
181..3054
/note="unnamed protein product; dystrophin related protein 2 (MGD|GI:107432, GB|U43520, evidence: BLASTN, 99%, match=807)
putative"
/codon_start=1
/protein_id="BAC38217.1"
/db_xref="GI:26349155"
/translation="MQPLVMQCPYTLPRCHEWHAADRFHHSSSLRNTCPQPVRAAV
TIPAPPDAGDPCLSPKLLNGTVGATGPLEPSAMNLCTWNEIKKSHNLRAPLEAFSD
LSGKLQPLREIIDWLSQKDELSAQLPLQGDVALVQOEKETHAFAFVEVKSQGPYIS
SVLESQAFLRSQHPFEELESQSEKSDTSRQRIQNLRFVWKQATVASSELWEKLTA
CVDQHRHEHTLEHLEIQGAMFELSSTLTQAEIGVFATWEPIDGLFIDSLPEHIQAIK
LFKEEFPVKDGVKLVNDLAHQIAISDVHLSMENSRALEQINIRWKQVSVAEERLKQ
LQAHDFGPGSOHFLSTSVQVPERAFSPNKVPYVYINHQAOITCWDHPKMTLYQTL
ADLNNTKFSAYRTAMKLRVQKALRLDLVLTLTALEIFNEHDLQASEHVMVDVEVIHC
LTALYERLEEEERGILNVPLCVDMSLNLLNVFDSRSGKMRALSFKTGIACLCGTEV
KEKQLVLSQVANSQSCDQRLHGLLHEAIQVPRQLGEVAFAFGSNVPSVRSFRR
STGKPVIEASQFLEWVNLEPQSMVWLAVLHRTVIAEQVKHOTKCSICROCPKGFYR
SLKQFNVDICQTCFLTGRASKGNKLHYPIMEYPTTSSNMRFDEATLLKNFRSKQY
FSKHPRGYLPVQSVLESDCSETPASSPMLPHADTHSRIEHFASRLAEMSONCSFFN
DLSPPDSIDEDQYLLRHSSPITDREPAFGQAPCSNATESKGELEKILAHLEDENRI
LQGEHLRLKWQHEAAEAPTLVEGSAEATPDHRNEELAEARILRQHSRLRTRMQIL
EDHNKQLESQRLRELQLPPSESDNGSAGSSLASPPQSEGHPRKQGTTPDTE
VADDVGSQSDVSLCLEDIMEKLRHAFPSVRSDDVTANTLLAS"

CDS

756 ACCGCTGGGTCTTTTACAAGACAGTTCTGACCAGTGAAGCGTCTGCACCTTCTCTGC 815
757 |||||
455 ACCTCCGCGCTCGCTAGAGGCCTTCTCAGACCTCAGTGGAAACCTCAGCTCCGCTCC 514
456 |||||
816 AGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTG 875
817 |||||
515 GAGAGATTATTGACTGGCTCAGCCAAAGGATGAGGAGTTGTGAGCTCAGCTCCCTTGC 574
516 |||||
876 GAGGCGACTTTCCAGCAGTTTCAAGACGACGATGATGATGATGATGATGATGATGAT 935
877 |||||
575 AAGGGATGTGGCCCTGGTACAACAGGAGAGGAGACACATGCAGCCCTTCATGGAAG 634
576 |||||
936 TGAATACTAAAGAACCTGTAATCATCATGAGTACTTCTGAGACTGTACGATATTTCTGAC 995
937 |||||
635 TCAAGTCTAAGGGCCCTACATCTCTCTCTGCTTGAATCCGCCAGGCTTTTCTGTCCC 694
636 |||||
996 AGCAGCCTTTGGAAGGACTAGAGAACTCTACCAGGAGCCAGAGAGCTGCCTCCTGAGG 1055
997 |||||
695 AGCATCCATTTGAAGAATTAGAGGAGTCTCAATCTGAGAGTAAAGATACTCTCCCCAG 754
1056 AGAGAGCCAGAAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGT 1115
1057 |||||
755 AGCGGATTCAAACCTTAGTCGCTTTTGTATGGAAGCAGGCAACAGTGGCCAGTGAGCTGT 814
756 |||||
1116 GGGAAAAATTGAACCTGCACTCCGCTGATCGGAGAGAAATTCGCTCTGAAAGAGAACGTGAGCC 1175
1117 |||||
815 GGGAGAAAGCTGACAGCCCGCTGTGTAGATCAGCATCGCCACATTGAGCATACTCTTGAAC 874
816 |||||
1176 GACTCCAGGAACCTCAAGAGGCCACCGGATGAGCTGAGCTCAAGCTGCGCCAAAGCTGAGG 1235
875 ATCTATTGGAGATCCAAGGGGCAATGGAGGAATTGAGCAGTACTTTGACCCCAAGCAGAGG 934
1236 TGATCAAGGATCTCTGGCAGCCCGCTGGGGGATCTCTCTCAATTGACTCTCTCCAAGATCACC 1295
935 GAGTCCGAGCCAGCTGGGAGCCCATAGGAGATCTCTTTATTGATTCTCCTCCAGAGCATA 994
1296 TCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCTCTGAAAGAGAACGTGAGCCACG 1355
995 TCCAAGCCATCAAGCTATTCAAAGAGAAATTCCTCTCTGTGAAAGATGGGTGAAGTTAG 1054
1356 TCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCA 1415
1055 TGAATGATCTGCCACCAGCTTGCCATCTCTGATGTGCACCTTGCAATGGAGAAATCCA 1114
1416 GCACTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCGCTCGAGGACCGAG 1475
1115 GGGCTCTGGAACAGATCAACATCCGGTGGAAACAGCTCCAGGTGTGAGTGGTGAGAGGC 1174
1476 TCAGGCAGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGCAGCTTTCTTTTCCA 1535
1175 TTAAGCAACTCCAGGATGCCCAACCGGACCTTTGGGCTGGGTCAAGCAGCTTCTCTCTCCA 1234
1536 CGTCTGTCCAGGTCCTTGGGAGAGAGCCATCTCGCCCAACAAAGTGCCTTACTATATCA 1595
1235 CTTCTGTCCAAGTTCCCTGGGAAAGAGCATTTTCTCCCAATAAAGTCCCTTACTACATCA 1294
1596 ACCACGAGACTCAAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTT 1655
1295 ACCACAGGCTCAGACCAACATGTCTGGGACCATCTTAAGATGACTGAGTTATACCAACCC 1354
1656 TAGCTGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCATGCAATGAAACTCCGAAGAC 1715
1355 TAGCCGATCTGAACAAACATTAAAGTTCTCAGCTTACCGCACTGCCATGGAAGCTCCGCAGAG 1414
1716 TGCAGAAAGGCCCTTGTGGATCTCTTGGAGCCTGTGAGCTGATGTGATGCTTGGACC 1775
1415 TCCAGAAAGGCCCTCGCCTGGATCTGGTAACTTAACTACAGCTCTGGAGATCTTCAATG 1474
1776 AGCACAACTCAAGCAAAATGACCAGCCCATGGATATCTCTGAGATTATTAATTTGTTGA 1835
1475 AGCATGACTTGCAGGCGCAGTGAACATGTGATGGATGTGGTGGAGGTCAATTCAGTCTTGA 1534

polyA_signal

3731..3736

/note="putative"

polyA_site

3753

/note="putative"

BASE COUNT
ORIGIN

927 a 1021 c 938 g 867 t

Query Match

24.2%; Score 509.2; DB 11; Length 3753;

Best Local Similarity 61.1%; Pred. No. 6.3e-96;

Matches 823; Conservative 0; Mismatches 523; Indels 0; Gaps 0;


```
source
1. .824
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DX001YB17"
/cell_line="HELA"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT      211 a   210 c   206 g   197 t
ORIGIN
Query Match      24.0%; Score 505; DB 9; Length 824;
Best Local Similarity 100.0%; Pred. No. 4e-95;
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2y 1597 CCACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTT 1656
Db 195 CCACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTT 254

2y 1657 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAATCCGAAGACT 1716
Db 255 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAATCCGAAGACT 314

2y 1717 GCAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGTGCATGTGATGCCCTTGGACCA 1776
Db 315 GCAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGTGCATGTGATGCCCTTGGACCA 374

2y 1777 GCACAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTTGAC 1836
Db 375 GCACAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTTGAC 434

2y 1837 CACTATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCCTCTCTGCGT 1896
Db 435 CACTATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCCTCTCTGCGT 494

2y 1897 GGATATGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 1956
Db 495 GGATATGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 554

2y 1957 TGTCCTGTCTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAGACAAGTA 2016
Db 555 TGTCCTGTCTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAGACAAGTA 614

2y 2017 CAGATACCTTTTCAAGCAAGTGCGAAGTTCAACAGGAATTTGTGACCAGCGCAGGCTGGG 2076
Db 615 CAGATACCTTTTCAAGCAAGTGCGAAGTTCAACAGGAATTTGTGACCAGCGCAGGCTGGG 674

2y 2077 CCTCCTCTGTCATGATTCATATCCAA 2101
Db 675 CCTCCTCTGTCATGATTCATATCCAA 699

RESULT 7
CB991394      801 bp      mRNA      linear      EST 01-MAY-2003
LOCUS
DEFINITION
AGENCOURT_13627932 NIH_MGC_148 Homo sapiens cDNA clone
IMAGE:30336570 5', mRNA sequence.

ACCESSION
CB991394      GI:30285818
VERSION
CB991394.1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 801)
NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
```

```
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM360 row: d column: 19
High quality sequence stop: 621.
Location/Qualifiers
1. .801
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30336570"
/tissue_type="pre-eclamptic placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH MGC 148"
/note="Organ: placenta; Vector: pBluescriptR; Site_1:
all-XhoI; Site_2: BamH; Library is oligo-dT primed and
directionally cloned using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert
size 2.3 kb and normalized to ROT 5. This is a primary
library enriched for full-length clones and constructed
using the Cap-trapper method (Carninci, in preparation).
Library constructed by M. Brownstein (NHGRI,
National Institutes of Health). Note: this is a NIH_MGC
Library."

BASE COUNT      204 a   210 c   199 g   188 t
ORIGIN
Query Match      24.0%; Score 503.4; DB 14; Length 801;
Best Local Similarity 99.8%; Pred. No. 8.7e-95;
Matches 504; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1597 CCACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTT 1656
Db 140 CCACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTT 199

Qy 1657 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAATCCGAAGACT 1716
Db 200 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAATCCGAAGACT 259

Qy 1717 GCAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGGACCA 1776
Db 260 GCAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGGACCA 319

Qy 1777 GCACAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTTGAC 1836
Db 320 GCACAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTTGAC 379

Qy 1837 CACTATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCCTCTCTGCGT 1896
Db 380 CACTATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCCTCTCTGCGT 439

Qy 1897 GGATATGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 1956
Db 440 GGATATGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 499

Qy 1957 TGTCCTGTCTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAGACAAGTA 2016
Db 500 TGTCCTGTCTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAGACAAGTA 559

Qy 2017 CAGATACCTTTTCAAGCAAGTGCGAAGTTCAACAGGAATTTGTGACCAGCGCAGGCTGGG 2076
Db 560 CAGATACCTTTTCAAGCAAGTGCGAAGTTCAACAGGAATTTGTGACCAGCGCAGGCTGGG 619

Qy 2077 CCTCCTCTGTCATGATTCATATCCAA 2101
Db 620 CCTCCTCTGTCATGATTCATATCCAA 644
```

RESULT 8
BX365572 1098 bp mRNA linear EST 05-MAY-2003
LOCUS
DEFINITION BX365572 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
CDNA clone CSODK001YB21 5-PRIME, mRNA sequence.
ACCESSION BX365572
VERSION BX365572.1 GI:30366927
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1098)
AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1955.r
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSIAK001ZB11QP1.
Location/Qualifiers
1. .1098
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODK001YB21"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 286 a 271 c 268 g 262 t 11 others
ORIGIN
Query Match 24.0%; Score 503.4; DB 13; Length 1098;
Best Local Similarity 99.8%; Pred. No. 9e-95;
Matches 504; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1597 CCACGAGACTCAAACTTGTCTGGGACCATCCAAATGACAGAGCTCTACCGTCTTT 1656
Db 204 CCACGAGACTCAAACTTGTCTGGGACCATCCAAATGACAGAGCTCTACCGTCTTT 263
QY 1657 AGCTGACCTGAATAATGTCTAGATCTCAGCTTATAGGACTGCCATGAACTCCGAGACT 1716
Db 264 AGCTGACCTGAATAATGTCTAGATCTCAGCTTATAGGACTGCCATGAACTCCGAGACT 323
QY 1717 GCAGAAGGCCCTTTGGTGGATCTCTTGAGCTGTCTAGCTGCATGTGATGCCCTTGACCA 1776
Db 324 GCAGAAGGCCCTTTGGTGGATCTCTTGAGCTGTCTAGCTGCATGTGATGCCCTTGACCA 383
QY 1777 GCACAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAAATGTTTGAC 1836
Db 384 GCACAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAAATGTTTGAC 443
QY 1837 CACTATTATGACCGCTGGAGCAAGAGCAGCAACAATTTGGTCAAGCTCCCTCTCTGCGT 1896
Db 444 CACTATTATGACCGCTGGAGCAAGAGCAGCAACAATTTGGTCAAGCTCCCTCTCTGCGT 503
QY 1897 GGATATGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCG 1956
Db 504 GGATATGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCG 563
QY 1957 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA 2016
Db 564 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA 623

QY 2017 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGG 2076
Db 624 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGG 683
QY 2077 CCTCCTTCTGCATGATTCTATCCAA 2101
Db 684 CCTCCTTCTGCATGATTCTATCCAA 708
RESULT 9
BQ640063 620 bp mRNA linear EST 15-JUL-2002
LOCUS
DEFINITION BQ640063 he23g04.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he
Homo sapiens cDNA clone he23g04 5', mRNA sequence.
ACCESSION BQ640063
VERSION BQ640063.1 GI:21764522
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 620)
AUTHORS Wistow,G., Bernstein,S.L., Wyatt,M.K., Ray,S., Behal,A., Touchman
J.W., Bouffard,G., Smith,D. and Peterson,K.
TITLE Expressed sequence tag analysis of human retina for the NEIBank
Project: Retbindin, an abundant, novel retinal cDNA and alternative
splicing of other retina-preferred gene transcripts
Mol. Vis. 8 (4), (2002) In press
JOURNAL Contact: Wistow G
COMMENT Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 23 row: g column: 04
Seq primer: M13RP1 reverse primer (ABI).
Location/Qualifiers
1. .620
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="he23g04"
/tissue_type="Retina"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Human Retina cDNA (Un-normalized, unamplified
): hd/he"
/note="Organ: Eye; Vector: pSPORT1; Neural retina tissue
was dissected from two 80 year old donors with no observed
eye disease. 100ug of total RNA was used for library
construction. A directionally cloned cDNA library in the
pSPORT1 vector (Life Technologies) was constructed at
Bioserve Biotechnology (Laurel MD) essentially following
the protocols of the SuperScript Plasmid System full
details of which are contained in the manufacturer's
instruction manual (http://www.lifetech.com/). First
strand synthesis was carried out using a Not I
primer-adaptor [5'-pGACTAGTTCTAGATCGGCGGCCGCC(T)15-3'
]. EST analysis was performed on the unamplified library
at the NIH Intramural Sequencing Center (NISC)."
BASE COUNT 165 a 148 c 144 g 163 t
ORIGIN
Query Match 23.5%; Score 494; DB 13; Length 620;
Best Local Similarity 100.0%; Pred. No. 7.8e-93;
Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1608 AAACAACCTTGTGGACCATCCAAATGACAGAGCTCTACCGTCTTTAGCTGACCTGA 1667
Db 1 AAACAACCTTGTGGACCATCCAAATGACAGAGCTCTACCGTCTTTAGCTGACCTGA 60

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM3135 row: e column: 11
High quality sequence stop: 583.
Location/Qualifiers

FEATURES

source
1. .728
/organism="Macaca mulatta"
/mol_type="mRNA"
/db_xref="taxon:9544"
/clone="IMAGE:6884820"
/tissue_type="Ovary"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD_Rh_Ov1"
/note="Organ: ovary; Vector: pDNR-LIB; Site 1: Sfi I; Site 2: Sfi I; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.0-4.0 kb. Tissue pooled from pre-pubertal, post pubertal sn menopausal monkeys. Constructed by Clontech. Note: this is a NICHD Library."
256 a 127 c 162 g 182 t 1 others

BASE COUNT
ORIGIN

Query Match 22.8%; Score 480; DB 14; Length 728;
Best Local Similarity 90.3%; Pred. No. 6.7e-90;
Matches 513; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 295 GGTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGAGAAATATTCAGAGA 354
DB 1 GGTGGTAATATTCTACAATTGGGAAGTCAGCTGATTGGAACAGGGAATATTCAGAGA 60
QY 355 TGAAGAACTGAAGTACAGAGCAGATGATCTCTAAATTCAGATGGGAATGCCTCAG 414
DB 61 TGAAGAACTGAAGTACAGAGCAGATGATCTCTAAATTCAGATGGGAATGCCTCAG 120
QY 415 GGTAGCTAGCATGGGAAACAAAGCAATTACATAGAGTTTTTAATGGATCTCCAGATCA 474
DB 121 GGTAGCTAGCATGGGAAACAAAGCAATTACATAGAGTTTTTAATGGATCTCCAGATCA 180
QY 475 GAAACTGAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAAGAACAAAGGAAATGGA 534
DB 181 GAAACTGAAAGAGTTGAATGACTGGCTGACAAAACAGAGAAAGAACAAAGGAAATGGA 240
QY 535 GGAAGAGCCTCTTTGACCTGTATCTTGAAGACCTAAACGCCAAGTACAAACATAAGGT 594
DB 241 GAAAGAACCCCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTACAAACATAAGGT 300
QY 595 GCTTCAAGAGAGATCTAGAACAAAGAACAAAGTCAAGGTCAATCTCTCACTCACATGGTGT 654
DB 301 GCTTCAAGAGAGATCTAGAACAAAGAACAAAGTCAAGGTCAATCTCTCACTCACATGGTGT 360
QY 655 GGTAGTTGATGAATCTAGTGGAGATCACGCCAACTGCTGCTTTGGAAGAACAACTTAAGGT 714
DB 361 GGTAGTTGATGAATCTAGTGGAGATCACGCCAACTGCTGCTTTGGAAGAGCAACTTAAGGT 420
QY 715 ATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACA 774
DB 421 ATTGGGAGATCGATGGGCAAAACATCTGCAGATGGACAGAGACCGCTGGGTTCTTTTACA 480
QY 775 AGACAGTTCTGACCACTGGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGGCT 834
DB 481 AGACATCCTTCTCAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCT 540
QY 835 ACAGCTGAAAGATGATGAATTAAGCCGG 862
DB 541 TTCAGAAAAGAGATGTGGTGAACAGG 568

RESULT 12
BG719710
LOCUS
DEFINITION 602690430F1 NIH_MGC_97 Homo sapiens cdna clone IMAGE:4822807 5',
mRNA sequence.
770 bp mRNA linear EST 08-MAY-2001
BG719710
602690430F1 NIH_MGC_97 Homo sapiens cdna clone IMAGE:4822807 5',
mRNA sequence.

ACCESSION

VERSION BG719710.1 GI:13998897

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 770)

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1CM10731 row: 1 column: 08

High quality sequence stop: 767.

FEATURES

source

Location/Qualifiers

1. .770

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4822807"

/lab_host="DH10B"

/clone_lib="NIH_MGC_97"

/note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: Sali-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."
189 a 199 c 198 g 184 t

BASE COUNT

ORIGIN

Query Match 22.4%; Score 471.4; DB 10; Length 770;
Best Local Similarity 98.2%; Pred. No. 4.2e-88;
Matches 498; Conservative 0; Mismatches 6; Indels 3; Gaps 2;

QY 1597 CCACGAGACTCAAACTTGTGGGACCATCCAAATGACAGAGCTCTACCAGTCTTT 1656
DB 150 CCACGAGACTCAAACTTGTGGGACCATCCAAATGACAGAGCTCTACCAGTCTTT 209
QY 1657 AGCTGACCTGAATAATGTCTGAGTTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACT 1716
DB 210 AGCTGACCTGAATAATGTCTGAGTTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACT 269
QY 1717 GCAGAAGGCCCTTGTGGATCTCTTGAGCCTGTCTGAGCTGATGCTTGGACCA 1776
DB 270 GCAGAAGGCCCTTGTGGATCTCTTGAGCCTGTCTGAGCTGATGCTTGGACCA 329
QY 1777 GCACAA--CCTCAAGCAAAATGACGAGCCCATGGATATCCTGAGATTATTATTGTTG 1834
DB 330 GCACAAATCCTCAAGTCAGAATGACGAGCCCATGGATATCCTGAGATTATTATTGTTG 389
QY 1835 ACCACTATTTATGACCGCCTGGAGCAAGAGCACAAATTTGTTCAACGTCCTCTCTGC 1894
DB 390 ACCACTATTTATGACCGCCTGGAGCAAGAGCACAAATTTGTTCAACGTCCTCTCTGC 449
QY 1895 GTGGATATGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATC 1954
DB 450 GTGGATATGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATC 509
QY 1955 CGTGTCTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAG 2014

Db 510 CGTGTCCTGCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAACAAG 569

QY 2015 TACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTG 2074

Db 570 TACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGA-TTTGTGACCAGCGCAGGCTG 628

QY 2075 GGCCTCCTTCTGTCATGATTCTATCCAA 2101

Db 629 GGCCTCCTTCTGTCATGATTCTATCCAA 655

RESULT 13

HSM083521

ID HSM083521 standard; RNA; EST; 591 BP.

XX AC EX490860;

XX SV BX490860.1

DT 09-MAY-2003 (Rel. 75, Created)

DT 09-MAY-2003 (Rel. 75, Last updated, Version 1)

XX Homo sapiens mRNA; EST DKFZp686F1392_r1 (from clone DKFZp686F1392)

KW EST; expressed sequence tag.

XX Homo sapiens (human)

DS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;

DC Eutheria; Primates; Catarrhini; Hominidae; Homo.

CX [1]

ZN 1-591

RA Koehrer K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,

RA Han M., Wiemann S.;

RT ;

RL Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases.

RL MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY

CX This is the 5' sequence of the clone insert

CC Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

CC Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

CC sequenced by BmFZ (Biomedical Research Center at the

CC Heinrich-Heine-University, Duesseldorf/Germany) within the cDNA

CC sequencing consortium of the German Genome Project.

CC No s1 sequence available.

CC This clone (DKFZp686F1392) is available at the RZPD in Berlin.

CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,

CC 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de

XX Key Location/Qualifiers

FH source

FT 1..591

FT /db_xref="taxon:9606"

FT /mol_type="mRNA"

FT /organism="Homo sapiens"

FT /clone="DKFZp686F1392"

FT /clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host

FT DH10B; sites SfiIA + SfiIB"

FT /dev_stage="adult"

FT /tissue_type="cDNA-collection"

CX Sequence 591 BP; 148 A; 157 C; 148 G; 137 T; 1 other;

IQ Query Match 21.7%; Score 455.8; DB 2; Length 591;

Best Local Similarity 99.3%; Pred. No. 7.5e-85;

Matches 457; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

DY 1597 CCACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAATGACAGAGCTCTACAGTCTTT 1656

Db 132 CCACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAATGACAGAGCTCTACAGTCTTT 191

DY 1657 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 1716

Db 192 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 251

QY 1717 GCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCA 1776

Db 252 GCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCA 311

QY 1777 GCACAACCTCAAGCAAAATGACAGCCCCATGGATATCCTGCAGATTATTAATTGTTTGAC 1836

Db 312 GCACAACCTCAAGCAAAATGACAGCCCCATGGATATCCTGCAGATTATTAATTGTTTGAC 371

QY 1837 CACTATTATGACCGCCTGGAGCAAGAGCACAAATTTGTTCAACGTCCTCTCTGCGT 1896

Db 372 CACTATTATGACCGCCTGGAGCAAGAGCACAAATTTGTTCAACGTCCTCTCTGCGT 431

QY 1897 GGATATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 1956

Db 432 GGATATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 491

QY 1957 TGTCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA 2016

Db 492 TGTCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA 551

QY 2017 CAGATACCTTTTCAAGCAAGTGCGCAAGTTTCAACAGGATTT 2056

Db 552 CAGATACCTTTTCAAGCANGTGCGCAAGTTTCAACAGGGATT 591

BB629984 652 bp mRNA linear EST 31-AUG-2001

BB629984 RIKEN full-length enriched, adult female vagina Mus

musculus cDNA clone 9930028B14 5', mRNA sequence.

BB629984

BB629984.1 GI:15399733

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 652)

Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,

, Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda

, M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,

Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki

, D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,

Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,

Muramatsu,M. and Hayashizaki,Y.

RIKEN Mouse ESTs (Arakawa,T., et al. 2001)

Unpublished

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.go.jp,

URL:http://genome.gsc.riken.go.jp/

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh

, M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,

Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura

, S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and

Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara

, Y. and Hayashizaki,Y.

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Yamanaka,I., Kiyosawa,H., Kondo,S., Saito,T., Shinagawa,A., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K., Arakawa,T., Ishii,Y. and Hayashizaki,Y.
Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct. Func. Genomics 2 pre, L72-L86 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
e mouse tissues.

FEATURES
source

Location/Qualifiers
1..652
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="9930028B14"
/sex="female"
/tissue_type="vagina"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult female vagina"
/note="Site_1: Sail; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'
GAGAGAGAGAAGGATCCAAAGAGCTCTTTTCTTTTCTTTTCTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTATCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 172 a 160 c 148 g 172 t
ORIGIN

Query Match 21.6%; Score 453.8; DB 10; Length 652;
Best Local Similarity 93.7%; Pred. No. 2e-84;
Matches 473; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 1597 CCACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAATGACAGAGCTCTACCACTTTT 1656
Db 45 CCACGAGACCCAAACCACTTGTGGGACCACCCCAAATGACAGAGCTCTACCACTTTT 104
QY 1657 AGCTGACCTGAATAATGTGATTTCTCAGATTCAGACTTATAGGACTGCATGAACCTCGAAGACT 1716
Db 105 AGCTGACCTGAATAATGTGAGTTCTCCGCGTATAGGACTGCCATGAAGCTCAGAAGCT 164
QY 1717 GCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGATGTGATGCCCTTGACCA 1776
Db 165 CCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGATGTGATGCCCTTGACCA 224
QY 1777 GCACAACTCAAGCAAAATGACCAGCCCATGGATATCCTGCGATATTATTAATTGTTGAC 1836
Db 225 GCACAACTCAAGCAAAATGACCAGCCCATGGATATCCTGCGATATTATTAATTGTTGAC 284
QY 1837 CACTATTTATGACCGCCTTGAGCAAGAGCACACAATTGTTGTTCAACGTCCTCTCTCGGT 1896
Db 285 TACAATTTATGATCGTCTGGAGCAAGAGCACACAATCTGTTCAATGTCCTCTCTGTGT 344
QY 1897 GGATATGTTGTAAGTGGTCTGATGTTTATGATACGGGACCAACAGGGAGGATCCG 1956
Db 345 GGATATGTTGTTCAACTGGCTTCTCAATGTTTATGATACGGGACCAACAGGGAGGATCCG 404
QY 1957 TGTCTGTCTTTTAAACTGGCATCATTCCTGTGTAAAGCACATTTGGAAGACAAGTA 2016
Db 405 TGTCTGTCTTTTAAACTGGCATCATTCCTGTGTAAAGCACACTTGAAGACAAGTA 464

QY 2017 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGCAGGCTGG 2076
Db 465 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACTGGCTTTGTGACCAGCGTAGGCTGG 524
QY 2077 CCTCCTTCTGTCATGATTCTATCCAA 2101
Db 525 TCTTCTTCTGTCATGATTCTATCCAA 549

RESULT 15
AK087829
LOCUS
DEFINITION

AK087829 1298 bp mRNA linear HTC 05-DEC-2002
Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN
full-length enriched library, clone:E330026B12 product:dystrophin,
muscular dystrophy, full insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AK087829
AK087829.1 GI:26352777
HTC; CAP trapper.
Mus musculus (house mouse)

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE

1
Carninci,P. and Hayashizaki,Y.

AUTHORS
TITLE
JOURNAL
MEDLINE
PubMed

High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

REFERENCE

AUTHORS

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

REFERENCE

AUTHORS

Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

REFERENCE

AUTHORS

Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
Toyo-oka,K., Wang,K.H., Weitz,C., Whitaker,C., Wilming,L.,
Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawai,J., Kohsaki,S.
and Hayashizaki,Y.

TITLE
JOURNAL
MEDLINE

Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660

PUBMED 11217851
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 1298)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoch,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.

FEATURES
source
1..1298
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:E330026B12"
/db_xref="taxon:10090"
/clone="E330026B12"
/sex="female"
/tissue type="ovary"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days pregnant adult"
120..1142
/note="unnamed protein product; dystrophin, muscular dystrophy (MGD|MG1:94909, GB|NM_007868, evidence: BLASTN, 100%, match=999)
putative"
/codon_start=1
/protein_id="BAC40019.1"
/db_xref="GI:26352778"
/translation="MREHLKGHETQTTCWDHPKMTELYQSLADLNNVRESAYRTAMKL RRLQKALCLDLSAACDALDQHNLKQNDQPMDIQIINCLTTIYDRLEQEHNNLVN VPLQVDMCLNWLNNVYDTGTRIRVLSFKTGII SLCKAHLEDKYRLEFKQVASSTGF CDQRRLGLLLHDSIQIPRLGEVASFGGSNIEPSVSCFQFANNKPEIHAALFLDWMR LEQSMVWLPVLHRVAAETAKHQAKCNI CKECPIIGFRYRSLKHFNYDICQSCFFSG RYAKGHKMHYPMVEYCTPTTSGEDVRDFAKVLKNKFRTRKRYFAKHPRMGVLPVQTVLE GDNMET"

BASE COUNT 343 a 332 c 296 g 327 t
ORIGIN
Query Match 21.6%; Score 453.8; DB 11; Length 1298;
Best Local Similarity 93.7%; Pred. No. 2.1e-84;
Matches 473; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
QY 1597 CCACGAGACTCAAACTTGCTGGACCATCCCAAAATGACAGAGCTCTACCACTCTT 1656
Db 140 CCACGAGACCCAAACCACTTGTGTTGGACCAACCCCAAAATGACAGAGCTCTACCACTCTT 199

QY 1657 AGCTGACCTGAATAATGT CAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 1716
Db 200 AGCTGACCTGAATAATGT CAGGTTCTCCGCGTATAGGACTGCCATGAAGCTCAGAAGGCT 259
QY 1717 GCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGT CAGCTGCATGTGATGCCCTTGACCA 1776
Db 260 CCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGT CAGCTGCATGTGATGCCCTTGACCA 319
QY 1777 GCACAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGAC 1836
Db 320 GCACAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATAATTAAGTGTGAC 379
QY 1837 CACTATTATTATGACCGCCTGGAGCAAGAGACAACAATTTGGTCAACGTCCCTCTCTGCGT 1896
Db 380 TACAATTTATGATCGTCTGGAGCAAGAGACAACAATCTGGTCAATGTCCCTCTCTGTGT 439
QY 1897 GGATATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 1956
Db 440 GGATATGTGTCTCAACTGGCTTCTCAATGTTTATGATACGGGACGAACAGGGAGGATCCG 499
QY 1957 TGTCTGTCTTTTAAAACTGGCATCATATTTCCCTGTGTATAAGCAATTTGGAAGACAAGTA 2016
Db 500 TGTCTGTCTTTTAAAACTGGCATCATATTTCTGTGTATAAGCACACTTGGAAAGACAAGTA 559
QY 2017 CAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTACCCAGCGCAGGCTGGG 2076
Db 560 CAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACTGGCTTTTGTGACCCAGCGTAGGCTGGG 619
QY 2077 CCTCCTTCTGCATGATTCTATCCAA 2101
Db 620 TCTTCTTCTGCATGATTCTATTCAA 644

Search completed: February 2, 2004, 06:03:17
Job time : 3941.93 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 10:14:55 ; Search time 7128.21 Seconds
(without alignments)
11483.979 Million cell updates/sec

Title: US-09-845-416-14_COPY_1000_3000
Perfect score: 2001
Sequence: 1 ggcagtcattgatggagag.....cattgagccaagtgtccgga 2001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------------|--------------------|
| 1 | 1820.8 | 91.0 | 5417 | 6 AX538619 | AX538619 Sequence |
| 2 | 1332 | 66.6 | 4402 | 6 E30220 | E30220 Shortened d |
| 3 | 1269.2 | 63.4 | 4402 | 6 E30219 | E30219 Shortened d |
| 4 | 1254 | 62.7 | 5339 | 6 AX538620 | AX538620 Sequence |
| 5 | 1132.2 | 56.6 | 5462 | 6 AX538621 | AX538621 Sequence |
| 6 | 1002.6 | 50.1 | 5952 | 6 AR304538 | AR304538 Sequence |
| 7 | 1002.6 | 50.1 | 5952 | 6 AX114289 | AX114289 Sequence |
| 8 | 1002.6 | 50.1 | 8689 | 6 AX538622 | AX538622 Sequence |
| 9 | 1002.6 | 50.1 | 11443 | 6 AX538624 | AX538624 Sequence |
| 10 | 1002.6 | 50.1 | 12057 | 6 AX538627 | AX538627 Sequence |
| 11 | 1002.6 | 50.1 | 12446 | 9 HSDMDR | X14298 Human mRNA |
| 12 | 1002.6 | 50.1 | 13957 | 6 AX409637 | AX409637 Sequence |
| 13 | 1002.6 | 50.1 | 13957 | 6 AX538581 | AX538581 Sequence |
| 14 | 1002.6 | 50.1 | 13957 | 9 HUMDYS | M18533 Homo sapien |
| 15 | 991.6 | 49.6 | 13977 | 6 AR220819 | AR220819 Sequence |
| 16 | 885.8 | 44.3 | 13887 | 4 AF070485 | AF070485 Canis fam |
| 17 | 869.8 | 43.5 | 13815 | 6 AX306153 | AX306153 Sequence |
| 18 | 869.8 | 43.5 | 13815 | 6 AX538582 | AX538582 Sequence |
| 19 | 869.8 | 43.5 | 13815 | 10 MUSDYSA | M68859 Mouse dyltr |
| 20 | 869.8 | 43.5 | 19307 | 6 AR093392 | AR093392 Sequence |
| 21 | 869.8 | 43.5 | 19307 | 6 AR142592 | AR142592 Sequence |
| 22 | 776.6 | 38.8 | 3275 | 10 MUSDYS | M18025 Mouse dyltr |
| 23 | 735.4 | 36.8 | 4075 | 6 E30221 | E30221 Shortened d |
| 24 | 731.2 | 36.5 | 3747 | 6 E30218 | E30218 Shortened d |
| 25 | 722.6 | 36.1 | 13575 | 5 GGDYS | X13369 Chicken mRN |
| 26 | 677.6 | 33.9 | 3163 | 6 E30223 | E30223 Shortened d |
| 27 | 615.4 | 30.8 | 2654 | 5 FSCDYSTRO | M37645 Torpedo cal |
| 28 | 612.2 | 30.6 | 630 | 9 HSDMDFL | X06179 Human fetal |
| 29 | 569 | 28.4 | 2110 | 9 HUMDMDXX | M92650 Human Duche |
| 30 | 569 | 28.4 | 4658 | 9 BC028720 | BC028720 Homo sapi |
| 31 | 534 | 26.7 | 3521 | 5 AF339031 | AF339031 Danio rer |
| 32 | 528 | 26.4 | 1966 | 9 HSDMDAL | X06178 Human adult |
| 33 | 518 | 25.9 | 1240 | 10 RNAPDY3 | X69767 R.norvegicu |
| 34 | 514.6 | 25.7 | 10705 | 10 RNAJ2967 | AJ002967 Rattus no |
| 35 | 506.6 | 25.3 | 3161 | 10 MMGUTRPH | X83506 M.musculus |
| 36 | 506.6 | 25.3 | 11096 | 6 AX538584 | AX538584 Sequence |
| 37 | 506.6 | 25.3 | 11096 | 10 MMY12229 | Y12229 M.musculus |
| 38 | 490.6 | 24.5 | 6045 | 6 A63605 | A63605 Sequence 7 |
| 39 | 490.6 | 24.5 | 6045 | 6 AR281528 | AR281528 Sequence |
| 40 | 490.6 | 24.5 | 6059 | 6 AX107972 | AX107972 Sequence |
| 41 | 490.6 | 24.5 | 10302 | 6 AX538583 | AX538583 Sequence |
| 42 | 490.6 | 24.5 | 10302 | 9 HSMUPS | X69086 H.sapiens m |
| 43 | 490.6 | 24.5 | 10320 | 6 A63607 | A63607 Sequence 9 |
| 44 | 490.6 | 24.5 | 10320 | 6 AR281529 | AR281529 Sequence |
| 45 | 475 | 23.7 | 3073 | 10 AF195788 | AF195788 Rattus no |

ALIGNMENTS

| | | | | | |
|--|------------|------------------------------------|---------|-----|--------|
| RESULT 1 | | | | | |
| AX538619 | AX538619 | Sequence 39 from Patent WO0229056. | 5417 bp | DNA | linear |
| LOCUS | AX538619 | Sequence 39 from Patent WO0229056. | | | |
| DEFINITION | AX538619 | Sequence 39 from Patent WO0229056. | | | |
| ACCESSION | AX538619 | Sequence 39 from Patent WO0229056. | | | |
| VERSION | AX538619.1 | GI:25271163 | | | |
| KEYWORDS | | | | | |
| SOURCE | | | | | |
| ORGANISM | | | | | |
| REFERENCE | | | | | |
| AUTHORS | | | | | |
| TITLE | | | | | |
| JOURNAL | | | | | |
| 1 | | | | | |
| Chamberlain,J.S. and Harper,S.Q. | | | | | |
| Mini-dystrophin nucleic acid and peptide sequences | | | | | |
| Patent: WO 0229056-A 39 11-APR-2002; | | | | | |
| THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US) | | | | | |

| FEATURES | | Location/Qualifiers | |
|-----------------------|-----------------------------|--|---------------------|
| source | 1. .5417 | /organism="synthetic construct" | |
| | | /mol_type="genomic DNA" | |
| | | /db_xref="taxon:32630" | |
| | | /note="Synthetic" | |
| BASE COUNT | 1700 a 1192 c 1182 g 1343 t | | |
| ORIGIN | | | |
| Query Match | 91.0%; | Score 1820.8; | DB 6; Length 5417; |
| Best Local Similarity | 92.6%; | Pred. No. 0; | |
| Matches 1999; | Conservative 0; | Mismatches 2; | Indels 157; Gaps 2; |
| QY | 1 | GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAA | 60 |
| Db | 1199 | GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAA | 1258 |
| QY | 61 | GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCAAGGAGAGATTTCTAAT | 120 |
| Db | 1259 | GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCAAGGAGAGATTTCTAAT | 1318 |
| QY | 121 | GATGTGGAAGTGGTGAAGACCAGTTTCATACTCATGAGGGGTACATGATGGATTGACA | 180 |
| Db | 1319 | GATGTGGAAGTGGTGAAGACCAGTTTCTACTACTCATGAGGGGTACATGATGGATTGACA | 1378 |
| QY | 181 | GCCCATCAGGCGCGGTTGGTAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGA | 240 |
| Db | 1379 | GCCCATCAGGCGCGGTTGGTAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGA | 1438 |
| QY | 241 | AAATTATCAGAAAGATGAAGAAACTGAAGTACAAAGAGCAGATGAATCTCCTAAATTCAAGA | 300 |
| Db | 1439 | AAATTATCAGAAAGATGAAGAAACTGAAGTACAAAGAGCAGATGAATCTCCTAAATTCAAGA | 1498 |
| QY | 301 | TGGGAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG | 360 |
| Db | 1499 | TGGGAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG | 1558 |
| QY | 361 | GATCTCCAGAATC-GAAACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAGAAAGA | 419 |
| Db | 1559 | GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAGAAAGA | 1618 |
| QY | 420 | ACAAGGAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA | 479 |
| Db | 1619 | ACAAGGAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA | 1678 |
| QY | 480 | CAACAACATAAGTGCTTCAAGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTCTCTC | 539 |
| Db | 1679 | CAACAACATAAGTGCTTCAAGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTCTCTC | 1738 |
| QY | 540 | ACTCACATGGTGGTAGTGTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA | 599 |
| Db | 1739 | ACTCACATGGTGGTAGTGTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA | 1798 |
| QY | 600 | GAACAACTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGC | 659 |
| Db | 1799 | GAACAACTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGC | 1858 |
| QY | 660 | TGGGTTCTTTTACAAGACATCCTTCTCAATGGCAACGTCCTTACTGAAGAACAGTGCCTT | 719 |
| Db | 1859 | TGGGTTCTTTTACAAGACATCCTTCTCAATGGCAACGTCCTTACTGAAGAACAGTGCCTT | 1918 |
| QY | 720 | TTTAGTGCATGGCTTTCAGAAAAAGAGATGCAGTGAACAAGANTTCAACAACCTGGCTTT | 779 |
| Db | 1919 | TTTAGTGCATGGCTTTCAGAAAAAGAGATGCAGTGAACAAGANTTCAACAACCTGGCTTT | 1978 |
| QY | 780 | AAAGATCAAAATGAATGTTATCAAGTCTTCAAAAAACTGGCCGTTTAAAGCGGATCTA | 839 |
| Db | 1979 | AAAGATCAAAATGAATGTTATCAAGTCTTCAAAAAACTGGCCGTTTAAAGCGGATCTA | 2038 |
| QY | 840 | GAAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAGATCTTCTTTCAACA | 899 |
| Db | 2039 | GAAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAGATCTTCTTTCAACA | 2098 |

| | | | |
|----|------|--|------|
| QY | 900 | CTGAAGAATAAGTCAGTGACCCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGT | 959 |
| Db | 2099 | CTGAAGAATAAGTCAGTGACCCAGAAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGT | 2158 |
| QY | 960 | TGGGATAAATTTAGTCCAAAAACTTTGAAAAAGAGTACAGCACAGA----- | 1002 |
| Db | 2159 | TGGGATAAATTTAGTCCAAAAACTTTGAAAAAGAGTACAGCACAGATTTTCACAGGCTGTCAAC | 2218 |
| QY | 1003 | ----- | 1002 |
| Db | 2219 | ACCACTCAGCCATCACTAACACAGACAACTGTAATGGAACAGTAACACTACGGTGACCACA | 2278 |
| QY | 1003 | ----- | 1002 |
| Db | 2279 | AGGGAACAGATCCTGGTAAAGCATGCTCAAGAGGAACTTCCACCACCACCTCCCCAAAAG | 2338 |
| QY | 1003 | -----CCCTTGAAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAG | 1043 |
| Db | 2339 | AAGAGGCAGATTACTGTGGATCTTGAAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAG | 2398 |
| QY | 1044 | CTGGACCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGAT | 1103 |
| Db | 2399 | CTGGACCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGAT | 2458 |
| QY | 1104 | CTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGAAATT | 1163 |
| Db | 2459 | CTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCACCTTCGAGGAGAAATT | 2518 |
| QY | 1164 | GCGCCTCTGAAAAGAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACITTG | 1223 |
| Db | 2519 | GCGCCTCTGAAAAGAGAACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACITTG | 2578 |
| QY | 1224 | GGCATTCAAGTCTCACCGTATAACCTCAGCAGCTCTGGAAGACCTGAAACACAGATGGAAG | 1283 |
| Db | 2579 | GGCATTCAAGTCTCACCGTATAACCTCAGCAGCTCTGGAAGACCTGAAACACAGATGGAAG | 2638 |
| QY | 1284 | CTTCTGCAGTGGCCCGTGCAGGACCGAGTCAGGCAGCTGATGAAGCCACAGGGACTTT | 1343 |
| Db | 2639 | CTTCTGCAGTGGCCCGTGCAGGACCGAGTCAGGCAGCTGATGAAGCCACAGGGACTTT | 2698 |
| QY | 1344 | GGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCCTGGGAGAGAGCCATC | 1403 |
| Db | 2699 | GGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCCTGGGAGAGAGCCATC | 2758 |
| QY | 1404 | TCGCCAAAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAACAACTTGTGGGACCAT | 1463 |
| Db | 2759 | TCGCCAAAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAACAACTTGTGGGACCAT | 2818 |
| QY | 1464 | CCCAAAATGACAGAGCTCTACCCAGTCTTTAGCTGACCTGTAATAATGTTCAGATTCTCAGCT | 1523 |
| Db | 2819 | CCCAAAATGACAGAGCTCTACCCAGTCTTTAGCTGACCTGTAATAATGTTCAGATTCTCAGCT | 2878 |
| QY | 1524 | TATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGC | 1583 |
| Db | 2879 | TATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGC | 2938 |
| QY | 1584 | CTGTCAGCTGCATGTGATGCCCTTGGACCAGCACAACTCAAGCAAAAATGACCGCCCATG | 1643 |
| Db | 2939 | CTGTCAGCTGCATGTGATGCCCTTGGACCAGCACAACTCAAGCAAAAATGACCGCCCATG | 2998 |
| QY | 1644 | GATATCCTGCAGATTATTAATTGTTTGACCACCTATTATGACCGCCTGGAGCAAGAGCAC | 1703 |
| Db | 2999 | GATATCCTGCAGATTATTAATTGTTTGACCACCTATTATGACCGCCTGGAGCAAGAGCAC | 3058 |
| QY | 1704 | AACAATTTGGTCAACAGTCCCTCTCTGCGTGGATATGTGTCTGAACCTGGCTGCTGAATGTT | 1763 |
| Db | 3059 | AACAATTTGGTCAACAGTCCCTCTCTGCGTGGATATGTGTCTGAACCTGGCTGCTGAATGTT | 3118 |
| QY | 1764 | TATGATACGGGACGAACAGGGAGGATCCGTGTCTCTTTTAAAACTGGCATCATTTCC | 1823 |
| Db | 3119 | TATGATACGGGACGAACAGGGAGGATCCGTGTCTCTTTTAAAACTGGCATCATTTCC | 3178 |
| QY | 1824 | CTGTGTAAAGCACATTTTGGAAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCA | 1883 |

| | | | | |
|------------|---|--|--|----------------------------|
| Db | 3179 | CTGTGTAAGCACATT | TTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCA | 3238 |
| Qy | 1884 | ACAGGATTTTGTGAC | CGCGCAGGTGGCCCTCCTTCTGCAATGATTTCTATCCAAATTTCCA | 1943 |
| Db | 3239 | ACAGGATTTTGTGAC | CGCGCAGGTGGCCCTCCTTCTGCAATGATTTCTATCCAAATTTCCA | 3298 |
| Qy | 1944 | AGACAGTTGGTGAAG | TTGCATCCTTTGGGGCAGTAACATTGAGCCAAAGTGTCCGGA | 2001 |
| Db | 3299 | AGACAGTTGGTGAAG | TTGCATCCTTTGGGGCAGTAACATTGAGCCAAAGTGTCCGGA | 3356 |
| RESULT 2 | | | | |
| E30220 | | | 4402 bp | DNA linear PAT 18-JUN-2001 |
| LOCUS | E30220 | | | |
| DEFINITION | Shortened dystrophin. | | | |
| ACCESSION | E30220 | | | |
| VERSION | E30220.1 | GI:13017027 | | |
| KEYWORDS | JP 1999318467-A/3. | | | |
| SOURCE | unidentified | | | |
| ORGANISM | unclassified. | | | |
| REFERENCE | 1 (bases 1 to 4402) | | | |
| AUTHORS | Sinichi,T. | | | |
| TITLE | Shortened dystrophin | | | |
| JOURNAL | Patent: JP 1999318467-A 3 24-NOV-1999; | | | |
| COMMENT | SCIENCE & TECH AGENCY,NATIONAL CENTER OF NEUROLOGY AND PSYCHIATRY | | | |
| | OS | Unidentified | | |
| | PN | JP 1999318467-A/3 | | |
| | PD | 24-NOV-1999 | | |
| | PF | 08-MAY-1998 JP 1998142134 | | |
| | PR | | | |
| | PI | SINICHI TAKEDA | | |
| | PC | C12N15/09,A61K48/00,C12N15/00 | | |
| | CC | Strandedness: Both; | | |
| | CC | Topology: Linear; | | |
| | FH | Key | Location/Qualifiers | |
| | FT | source | 1..4402 | |
| | FT | | /organism='Unidentified'. | |
| FEATURES | source | Location/Qualifiers | | |
| | | 1..4402 | | |
| | | /organism="unidentified" | | |
| | | /mol_type="genomic DNA" | | |
| | | /db_xref="taxon:32644" | | |
| BASE COUNT | 1339 a | 984 c | 1010 g | 1069 t |
| ORIGIN | | | | |
| | Query Match 66.6%; Score 1332; DB 6; Length 4402; | | | |
| | Best Local Similarity 83.8%; Pred.No. 0; | | | |
| | Matches 1677; Conservative 0; Mismatches 0; Indels 325; Gaps 2; | | | |
| Qy | 1 | GGCAGTTCAATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA | 60 | |
| Db | 1195 | GGCAGTTCAATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA | 1254 | |
| Qy | 61 | GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAAT | 120 | |
| Db | 1255 | GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAAT | 1314 | |
| Qy | 121 | GATGTGGAAGTGGTGAAGACCAGTTTCATACTCATGAGGGGTACATGATGGATTGACA | 180 | |
| Db | 1315 | GATGTGGAAGTGGTGAAGACCAGTTTCATACTCATGAGGGGTACATGATGGATTGACA | 1374 | |
| Qy | 181 | GCCCATCAGGCGCGGTTGGTAAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA | 240 | |
| Db | 1375 | GCCCATCAGGCGCGGTTGGTAAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA | 1434 | |
| Qy | 241 | AAATTATCAGAAGATGAAGAACTGAAGTACAAGACAGATGAATCTCCTAAATTCAAGA | 300 | |
| Db | 1435 | AAATTATCAGAAGATGAAGAACTGAAGTACAAGACAGATGAATCTCCTAAATTCAAGA | 1494 | |
| Qy | 301 | TGGGAATGCTCAGGGTAGCTAGCATGGAACAAAGCAATTTACATAGAGTTTAAATG | 360 | |

| | | | |
|----|------|--|------|
| Db | 1495 | TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGATTTTAAATG | 1554 |
| Qy | 361 | GATCTCCAGAAATC-GAAACTGAAAGAGTTGAATGACTGGTTAAACAAAAACAGAAGAAAGA | 419 |
| Db | 1555 | GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGTTAAACAAAAACAGAAGAAAGA | 1614 |
| Qy | 420 | ACAAGGAAATGGAGGAAGAGCCTCTTGGACCTTGATCTTGAAGACCTTAAACCGCCAAAGTA | 479 |
| Db | 1615 | ACAAGGAAATGGAGGAAGAGCCTCTTGGACCTTGATCTTGAAGACCTTAAACCGCCAAAGTA | 1674 |
| Qy | 480 | CAACAACATAAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTC | 539 |
| Db | 1675 | CAACAACATAAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTC | 1734 |
| Qy | 540 | ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAAACTGCTGCTTTGGAA | 599 |
| Db | 1735 | ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAAACTGCTGCTTTGGAA | 1794 |
| Qy | 600 | GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGC | 659 |
| Db | 1795 | GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGC | 1854 |
| Qy | 660 | TGGGTTCTTTTACAAGACATCCTTCTCAAATGGCAACGTTTACTGAAGAACAGTGCCTT | 719 |
| Db | 1855 | TGGGTTCTTTTACAAGACATCCTTCTCAAATGGCAACGTTTACTGAAGAACAGTGCCTT | 1914 |
| Qy | 720 | TTTAGTGCATGGCTTTTCAGAAAAAAGAGATGCAGTGAACAAGATTTCACACAACTGGCTTT | 779 |
| Db | 1915 | TTTAGTGCATGGCTTTTCAGAAAAAAGAGATGCAGTGAACAAGATTTCACACAACTGGCTTT | 1974 |
| Qy | 780 | AAAGATCAAAAATGAAATGTTTATCAAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTA | 839 |
| Db | 1975 | AAAGATCAAAAATGAAATGTTTATCAAAGT----- | 2001 |
| Qy | 840 | GAAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACA | 899 |
| Db | 2002 | ----- | 2001 |
| Qy | 900 | CTGAAGAATAAGTCAGTGACCCAGAAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGT | 959 |
| Db | 2002 | ----- | 2001 |
| Qy | 960 | TGGGATAAATTTAGTCCAAAAACTTGAAAAAGAGTACAGCACAGACCCCTTGAAAAGACTCCAG | 1019 |
| Db | 2002 | ----- | 2001 |
| Qy | 1020 | GAACTTCAAGAGGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAG | 1079 |
| Db | 2002 | ----- | 2001 |
| Qy | 1080 | GGATCCTGGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAA | 1139 |
| Db | 2002 | -----CTCGAGAAA | 2010 |
| Qy | 1140 | GTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCATGAC | 1199 |
| Db | 2011 | GTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCATGAC | 2070 |
| Qy | 1200 | CTTGCTGCCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCACTCTG | 1259 |
| Db | 2071 | CTTGCTGCCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCACTCTG | 2130 |
| Qy | 1260 | GAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGCAG | 1319 |
| Db | 2131 | GAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGCAG | 2190 |
| Qy | 1320 | CTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCGTCTC | 1379 |
| Db | 2191 | CTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCGTCTC | 2250 |
| Qy | 1380 | CAGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCACGAG | 1439 |
| Db | 2251 | CAGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCACGAG | 2310 |

Db 1863 -----AATAGATGAGACCCCTTGAAAGACTCCAG 1890
QY 1020 GAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGTGAGGTGATCAAG 1079
Db 1891 GAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGTGAGGTGATCAAG 1950
QY 1080 GGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGAGAAA 1139
Db 1951 GGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGAGAAA 2010
QY 1140 GTCAAGGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGAC 1199
Db 2011 GTCAAGGCACCTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGAC 2070
QY 1200 CTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCACTCTG 1259
Db 2071 CTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCACTCTG 2130
QY 1260 GAAGACCTGAACACACAGATGGAAGCTTTCGAGGTGGCCGTCGAGACCGAGTCAGGCAG 1319
Db 2131 GAAGACCTGAACACACAGATGGAAGCTTTCGAGGTGGCCGTCGAGACCGAGTCAGGCAG 2190
QY 1320 CTGCATGAAGCCACAGGGACTTTTGGTCAGCATCTCAGCACTTCTTTCCACGTCGTCTC 1379
Db 2191 CTGCATGAAGCCACAGGGACTTTTGGTCAGCATCTCAGCACTTCTTTCCACGTCGTCTC 2250
QY 1380 CAGGGTCCCTGGAGAGAGCCATCTCGCAAAACAAAGTGCCCTACTATATCAACCCAGAG 1439
Db 2251 CAGGGTCCCTGGAGAGAGCCATCTCGCAAAACAAAGTGCCCTACTATATCAACCCAGAG 2310
QY 1440 ACTCAAAACAACCTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGAC 1499
Db 2311 ACTCAAAACAACCTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGAC 2370
QY 1500 CTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAA 1559
Db 2371 CTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAA 2430
QY 1560 GCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCTCTTGACCAGCAAC 1619
Db 2431 GCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCTCTTGACCAGCAAC 2490
QY 1620 CTCAAGCAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACCTATT 1679
Db 2491 CTCAAGCAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACCTATT 2550
QY 1680 TATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCCTCTCTGCGTGGATATG 1739
Db 2551 TATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAACGTCCCTCTCTGCGTGGATATG 2610
QY 1740 TGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGCTCTG 1799
Db 2611 TGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGCTCTG 2670
QY 1800 TCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGATAC 1859
Db 2671 TCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGATAC 2730
QY 1860 CTTTTCAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCCAGCGCAGGCTGGGCCCTCCTT 1919
Db 2731 CTTTTCAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCCAGCGCAGGCTGGGCCCTCCTT 2790
QY 1920 CTGCATGATTCATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGCGAGT 1979
Db 2791 CTGCATGATTCATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGCGAGT 2850
QY 1980 AACATTGAGCCCAAGTGTCCGGA 2001
Db 2851 AACATTGAGCCCAAGTGTCCGGA 2872

RESULT 4

AX538620
LOCUS AX538620 5339 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 40 from Patent WO0229056.
ACCESSION AX538620
VERSION AX538620.1 GI:25271166
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.O.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 40 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
1. 5339
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"
BASE COUNT 1638 a 1191 c 1187 g 1323 t
ORIGIN

Query Match 62.7%; Score 1254; DB 6; Length 5339;
Best Local Similarity 77.9%; Pred. No. 4.9e-304;
Matches 1621; Conservative 0; Mismatches 380; Indels 79; Gaps 6;

QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA 60
Db 1199 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA 1258
QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCCTAAT 120
Db 1259 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCCTAAT 1318
QY 121 GATGTGGAAGTGGTGAAGACCAGTTTCATCTACTCATGAGGGGTACATGATGGATTTGACA 180
Db 1319 GATGTGGAAGTGGTGAAGACCAGTTTCATCTACTCATGAGGGGTACATGATGGATTTGACA 1378
QY 181 GCCCATCAGGGCCGGTGGTAAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240
Db 1379 GCCCATCAGGGCCGGTGGTAAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 1438
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGA 300
Db 1439 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGA 1498
QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 360
Db 1499 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATCATAGATTACTG 1558
QY 361 GATCTCCAGAATC-GAAACTGAAGAGATTGAATGACTGGCTAAACAAAAACAGAAAGA 419
Db 1559 CAACAGTTCCTCCCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGAAGCTGAACCAACT 1618
QY 420 ACAAGGAAAATGGAGGAAGAGCCCTCTTGACCTGATCTTGAAGACCTAAAAACGCCAAGTA 479
Db 1619 GCCAATGCTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTA 1678
QY 480 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAACAAAGTCAGGGTCAATTCCTCTC 539
Db 1679 AAAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGAATTTGAAGCTCACACAGATGTT 1738
QY 540 ACTCACATGGTGGTGGTAG-----TTGATGAATCTAGTGGAGATCACGCAACT 587
Db 1739 TATCACACCTGGATGAAAACACAGCCAAAAAATCCTTGAGATCCCTGGAAGGTTCCGATGAT 1798
QY 588 GCTGCTTTGGAAGAACAACTTAAG---TATTGGGAGATCGATGGGCAACATCTGTAGA 644
Db 1799 GCAGTCTCTTTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAGTGAACCTTCGAAA 1858
QY 645 TGGACAGAAGACCGCTGGGTTCTTTTACAAGACATCCTTCTCAAATGGCAACGTCCTTACT 704

Db 1859 AAGTCTCTCAACATTAGGTCCCATTTTGGAAAGCCAGTTCTGACAGTGGAAGCGTCTGCAC 1918
QY 705 GAAGAACAGTGCCTTTTATGTGATGCTTTTTCAGAAAAAGAAAGATGCAGTGAACAAGATT 764
Db 1919 CTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGAG 1978
QY 765 CACACAACCTGCTTTAAAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACTGSCCGTT 824
Db 1979 GCACCTATTGGAGGCGACTTTCAGCAGTTTCAAGAGCAGACGATGTACATAGGGCCTTC 2038
QY 825 TTAAAAAGCGATCTAGAAAAAGAAAAAGCAATCCATGGGCAAACTGTA----- 871
Db 2039 AAGAGGGAATTGAAGAACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATA 2098
QY 872 -TTCACCTCAACAAGATCTTCTTTCAACACTGAAGATAAGTCAGTGACCCAGAGACGG 930
Db 2099 TTTCTGACAGACGACCTTTTGAAGGACTAGAGAAACTCTACCAGGAGCCACAGAGCTG 2158
QY 931 AAGCATGGCTGGA-----TAACTTTGCCCGGTGTTGG 963
Db 2159 CCTCCTGAGGAGAGAGCCAGAAATGTCACTGGCTTCTACGAAAGCAGGCTGAGGAGGTC 2218
QY 964 ATAATTAGTCCAAAAAAGTTGAA-----AAGAGTACAGCACAG 1001
Db 2219 AATACTGAGTGGGAAAAAATTGAACTTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAG 2278
QY 1002 ACCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCGC 1061
Db 2279 ACCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCGC 2338
QY 1062 CAAGCTGAGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTC 1121
Db 2339 CAAGCTGAGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTC 2398
QY 1122 CAAGATCACCTCGAGAAAAAGTCAAGGCACTTCGAGGAGAAAAATTCGCGCTCTGAAAGAGAAC 1181
Db 2399 CAAGATCACCTCGAGAAAAAGTCAAGGCACTTCGAGGAGAAAAATTCGCGCTCTGAAAGAGAAC 2458
QY 1182 GTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGSCATTCAGCTCTCACCG 1241
Db 2459 GTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGSCATTCAGCTCTCACCG 2518
QY 1242 TATAACCTCAGCACTCTGGAAGACCTGAAACACCAAGATGGAAGCTTCTGCAGGTGGCCGTC 1301
Db 2519 TATAACCTCAGCACTCTGGAAGACCTGAAACACCAAGATGGAAGCTTCTGCAGGTGGCCGTC 2578
QY 1302 GAGGACCGAGTCAAGCAGCTGCATGAAGCCACAGGGACTTTGTGCCAGCATCTCAGCAC 1361
Db 2579 GAGGACCGAGTCAAGCAGCTGCATGAAGCCACAGGGACTTTGTGCCAGCATCTCAGCAC 2638
QY 1362 TTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCC 1421
Db 2639 TTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCC 2698
QY 1422 TACTATATCAACCAGAGACTCAAAACAATTTGCTGGGACCATCCCAAAATGACAGAGCTC 1481
Db 2699 TACTATATCAACCAGAGACTCAAAACAATTTGCTGGGACCATCCCAAAATGACAGAGCTC 2758
QY 1482 TACCAGTCTTTAGCTGACCTGAATAATGTGAGATTTCTCAGCTTATAGGACTGCCATGAAA 1541
Db 2759 TACCAGTCTTTAGCTGACCTGAATAATGTGAGATTTCTCAGCTTATAGGACTGCCATGAAA 2818
QY 1542 CTCGGAAGACTGCAGAGGCCCTTTTGGCTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGAT 1601
Db 2819 CTCGGAAGACTGCAGAGGCCCTTTTGGCTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGAT 2878
QY 1602 GCCTTGGACCGACCAACCTCAAGCAAATGACCGCCCATGGATATCCTGCAGATTATT 1661
Db 2879 GCCTTGGACCGACCAACCTCAAGCAAATGACCGCCCATGGATATCCTGCAGATTATT 2938
QY 1662 AATTGTTTGACCACTATTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTC 1721
Db 2939 AATTGTTTGACCACTATTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTC 2998

QY 1722 CCTCTCTGCGTGGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACA 1781
Db 2999 CCTCTCTGCGTGGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACA 3058
QY 1782 GGGAGGATCCGTGTCTCTCTTTTAAAACTGGCATCATTCCTCTGTGTAAAGCACATTG 1841
Db 3059 GGGAGGATCCGTGTCTCTCTTTTAAAACTGGCATCATTCCTCTGTGTAAAGCACATTG 3118
QY 1842 GAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAG 1901
Db 3119 GAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAG 3178
QY 1902 CGCAGGCTGGGCCCTCCTTCTGATGATTTCTATCCAAATCCAAAGACAGTTGGTGAAGTT 1961
Db 3179 CGCAGGCTGGGCCCTCCTTCTGATGATTTCTATCCAAATCCAAAGACAGTTGGTGAAGTT 3238
QY 1962 GCATCCTTTGGGGCGAGTAACATTCAGCCAAAGTGTCCGGA 2001
Db 3239 GCATCCTTTGGGGCGAGTAACATTCAGCCAAAGTGTCCGGA 3278

RESULT 5
AX538621
LOCUS AX538621 5462 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 41 from Patent WO0229056.
ACCESSION AX538621
VERSION AX538621.1 GI:25271168
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 41 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
source
1..5462
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"

BASE COUNT 1668 a 1225 c 1212 g 1357 t
ORIGIN

Query Match 56.6%; Score 1132.2; DB 6; Length 5462;
Best Local Similarity 73.9%; Pred. No. 2e-273;
Matches 1628; Conservative 0; Mismatches 373; Indels 202; Gaps 7;

QY 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAA 60
Db 1199 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAA 1258
QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAAT 120
Db 1259 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAAT 1318
QY 121 GATGTGGAAGTGGTGAAGACCAGTTTCATACTCATGAGGGGTACATGATGGATTGACA 180
Db 1319 GATGTGGAAGTGGTGAAGACCAGTTTCATACTCATGAGGGGTACATGATGGATTGACA 1378
QY 181 GCCCATCAGGCCCGGGTTGGTAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGA 240
Db 1379 GCCCATCAGGCCCGGGTTGGTAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGA 1438
QY 241 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 300
Db 1439 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 1498
QY 301 TGGGAATGCCTCAGGGTAGCTAGATGGAAAAACAAAGCAATTTACATAGATTTTAAATG 360
Db 1499 TGGGAATGCCTCAGGGTAGCTAGATGGAAAAACAAAGCAATTTACATGCTCCTGGACTG 1558

QY 361 ----- 360
Db 1559 ACCACTATTGGAGCCTCTCTCTACTCAGACTGTTACTCTGGTGACACAACTCTGGTTACT 1618
QY 361 -----GATCTCCAGAATCGAAA----- 377
Db 1619 AAGGAAACTGCCATCTCCAAACTAGAAAATGCCATCTCTCTTGATGTTGGAGCATAGATTA 1678
QY 378 -----CTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAGAA 416
Db 1679 CTGCAACAGTTCCCTCGACCTGGAAAAGTTCTTGCTGGCTTACAGAAGCTGAAACA 1738
QY 417 AGAACAAAGGAAATGGAGGAAGAGCCTCTTGACCTGATCTTGAGACCTTAAACGCCAA 476
Db 1739 ACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGA 1798
QY 477 GTACAAACACATAAGTGCTTCAAGAAGATCTAGAACAAGAACAACTCAGGGTCAATTCT 536
Db 1799 GTAAAAGAGCTGATGAAACAATGGCAAGACTCCAAGGTGAAATGAAGCTCACACAGAT 1858
QY 537 CTCACCTCACATGGTGGTGGTAGTTGATGA-----ATCTAGTGGAGATCACGCA 584
Db 1859 GTTTATCACAACTGGATGAACACAGCCCAAAAATCCTGAGATCCCTGGAAGGTTCCGAT 1918
QY 585 ACTGCTGCTTTGGAAGAACAACTTAAG---TATTGGGAGATCGATGGGCAACATCTGT 641
Db 1919 GATGCAGTCCCTGTACAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTCGG 1978
QY 642 AGATGGACAGAAACCGCTGGGTTCTTTTACAAGACATCCTTCTCAAAATGGCAACGCTTT 701
Db 1979 AAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTGAAGCGTGTG 2038
QY 702 ACTGAAGAACAGTGCCTTTTATGTGATGGCTTTTCAAAAAGAAAGATGCAGTGAAACAAG 761
Db 2039 CACCTTTCTCTGCAGGAACCTTCTGGTGGCTACAGCTGAAAGATGATGAATTAAGCCGG 2098
QY 762 ATTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACCTGGCC 821
Db 2099 CAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAAAGCAGAACAGATGTACATAGGGCC 2158
QY 822 GTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCATGGGCAAACTGTA----- 871
Db 2159 TTCAAGAGGGAATTGAAAACCTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGA 2218
QY 872 ----TTCACCTCAAACAAGATCTTCTTCAACACACTGAAGAAATAAGTCAAGTACCCAGAGA 927
Db 2219 ATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAACTCTACCAGGAGCCCAAGAGAG 2278
QY 928 CGGAAGCATGSGTGA-----TAACTTTGCCCGTGT 960
Db 2279 CTGCCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAG 2338
QY 961 GGGATAATTTAGTCCAAAACCTTGA-----AAGAGTACAGCA 998
Db 2339 GTCAATACTAGTGGGAAAATAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGAT 2398
QY 999 CAGACCCCTTGAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTG 1058
Db 2399 GAGACCCCTTGAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTG 2458
QY 1059 CGCCAAGCTGAGGTGATCAAGGATCTCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCT 1118
Db 2459 CGCCAAGCTGAGGTGATCAAGGATCTCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCT 2518
QY 1119 CTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAG 1178
Db 2519 CTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAG 2578
QY 1179 AACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACTTTGGGCATTCAGCTCTCA 1238
Db 2579 AACGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACTTTGGGCATTCAGCTCTCA 2638

QY 1239 CCGTATAAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCC 1298
Db 2639 CCGTATAAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCC 2698
QY 1299 GTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGACTTTGGTCCAGCATCTCAG 1358
Db 2699 GTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAAGGACTTTGGTCCAGCATCTCAG 2758
QY 1359 CACTTTCTTTCCAGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTG 1418
Db 2759 CACTTTCTTTCCAGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTG 2818
QY 1419 CCCTACTATATCAACACGAGACTCAAAACAACCTTGTGGGACCATCCCAAAATGACAGAG 1478
Db 2819 CCCTACTATATCAACACGAGACTCAAAACAACCTTGTGGGACCATCCCAAAATGACAGAG 2878
QY 1479 CTCTACCAGTCTTTAGCTGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATG 1538
Db 2879 CTCTACCAGTCTTTAGCTGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATG 2938
QY 1539 AAACCTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTTCAGCTGCATGT 1598
Db 2939 AAACCTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTTCAGCTGCATGT 2998
QY 1599 GATGCCCTTGGACGACACAACCTCAAGCAAAAATGACCAGCCCATGSGATATCTCTGCAGATT 1658
Db 2999 GATGCCCTTGGACGACACAACCTCAAGCAAAAATGACCAGCCCATGSGATATCTCTGCAGATT 3058
QY 1659 ATTAATTTGTTGACCACCTATTTATGACCCGCTGGAGCAAGAGCAACAATTTGGTCAAC 1718
Db 3059 ATTAATTTGTTGACCACCTATTTATGACCCGCTGGAGCAAGAGCAACAATTTGGTCAAC 3118
QY 1719 GTCCTCTCTCGTGGATATGTGTCTGAACTGGCTGCTGATGTTTATGATACGGGACGA 1778
Db 3119 GTCCTCTCTCGTGGATATGTGTCTGAACTGGCTGCTGATGTTTATGATACGGGACGA 3178
QY 1779 ACAGGGAGGATCCGTGTCCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACAT 1838
Db 3179 ACAGGGAGGATCCGTGTCCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACAT 3238
QY 1839 TTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGAC 1898
Db 3239 TTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGAC 3298
QY 1899 CAGCGCAGGCTGGCCCTCCTCTGATGATTTCTATCCAAATCCAAAGACAGTTGGTGAA 1958
Db 3299 CAGCGCAGGCTGGCCCTCCTCTGATGATTTCTATCCAAATCCAAAGACAGTTGGTGAA 3358
QY 1959 GTTGATCCTTTGGGGCAGTAACTTGAGCCAAGTGTCCGGA 2001
Db 3359 GTTGATCCTTTGGGGCAGTAACTTGAGCCAAGTGTCCGGA 3401

RESULT 6
AR304538
LOCUS AR304538 5952 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 1 from patent US 6544786.
ACCESSION AR304538
VERSION AR304538.1 GI:31693691
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5952)
AUTHORS Xiao,X. and Liu,P.X.
TITLE Method and vector for producing and transferring trans-spliced peptides
JOURNAL Patent: US 6544786-A 1 08-APR-2003;
FEATURES Location/Qualifiers
source 1. .5952
/organism="unknown"
BASE COUNT 1860 a 1344 c 1410 g 1338 t
ORIGIN

| | | | | | |
|-----------------------|------|---|---------------------|------------|----------------------|
| Query Match | | 50.1%; | Score 1002.6; | DB 6; | Length 5952; |
| Best Local Similarity | | 99.6%; | Pred. No. 7.8e-241; | | |
| Matches 1005; | | Conservative | 0; | Mismatches | 4; Indels 0; Gaps 0; |
| QY | 993 | ACAGCACAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTC | 1052 | | |
| Db | 3679 | ATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTC | 3738 | | |
| QY | 1053 | AAGCTGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCTCATT | 1112 | | |
| Db | 3739 | AAGCTGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCTCATT | 3798 | | |
| QY | 1113 | GACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCATCTCGAGGAGAAAATTGGCGCTCTG | 1172 | | |
| Db | 3799 | GACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCATCTCGAGGAGAAAATTGGCGCTCTG | 3858 | | |
| QY | 1173 | AAAGAGAACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAG | 1232 | | |
| Db | 3859 | AAAGAGAACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAG | 3918 | | |
| QY | 1233 | CTCTCACCGTATAACCTCAGCACCTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAG | 1292 | | |
| Db | 3919 | CTCTCACCGTATAACCTCAGCACCTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAG | 3978 | | |
| QY | 1293 | GTGGCCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCA | 1352 | | |
| Db | 4039 | GTGGCCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCA | 4098 | | |
| QY | 1353 | TCTCAGCATTTCTCTTCCAGCTCTGTCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAAC | 1412 | | |
| Db | 4039 | TCTCAGCATTTCTCTTCCAGCTCTGTCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAAC | 4098 | | |
| QY | 1413 | AAAGTGCCCTACTATATCAACCCAGAGACTCAAAACAACTTGTGGGAGAGCCATCCCAAATG | 1472 | | |
| Db | 4099 | AAAGTGCCCTACTATATCAACCCAGAGACTCAAAACAACTTGTGGGAGAGCCATCCCAAATG | 4158 | | |
| QY | 1473 | ACAGAGCTCTACAGACTCTTTAGTGTGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACT | 1532 | | |
| Db | 4159 | ACAGAGCTCTACAGACTCTTTAGTGTGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACT | 4218 | | |
| QY | 1533 | GCCATGAAACTCCGAAGACTGCAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCT | 1592 | | |
| Db | 4219 | GCCATGAAACTCCGAAGACTGCAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCT | 4278 | | |
| QY | 1593 | GCATGTGATGCCCTGGACCAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTG | 1652 | | |
| Db | 4279 | GCATGTGATGCCCTGGACCAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTG | 4338 | | |
| QY | 1653 | CAGATTATTAATTGTTTGACCACTATTATGACCGCCTGGAGCAAGAGCACAACAATTG | 1712 | | |
| Db | 4339 | CAGATTATTAATTGTTTGACCACTATTATGACCGCCTGGAGCAAGAGCACAACAATTG | 4398 | | |
| QY | 1713 | GTCAACGTCCTCTCTGCGTGGATATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACG | 1772 | | |
| Db | 4399 | GTCAACGTCCTCTCTGCGTGGATATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACG | 4458 | | |
| QY | 1773 | GGACGAACAGGGAGATCCGTGTCTTCTTTAAACCTGGCATCATTTCCCTGTGTAAA | 1832 | | |
| Db | 4459 | GGACGAACAGGGAGATCCGTGTCTTCTTTAAACCTGGCATCATTTCCCTGTGTAAA | 4518 | | |
| QY | 1833 | GCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTT | 1892 | | |
| Db | 4519 | GCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTT | 4578 | | |
| QY | 1893 | TGTGACCAAGCGCAGGTGGGCGCTCCTTCTGCATGATTCTATCCAAATTTCAAGACAGTTG | 1952 | | |
| Db | 4579 | TGTGACCAAGCGCAGGTGGGCGCTCCTTCTGCATGATTCTATCCAAATTTCAAGACAGTTG | 4638 | | |
| QY | 1953 | GGTGAAGTTGCATCCTTTGGGGGCGAGTAACATTGAGCCAAAGTGTCCGGA | 2001 | | |
| Db | 4639 | GGTGAAGTTGCATCCTTTGGGGGCGAGTAACATTGAGCCAAAGTGTCCGGA | 4687 | | |

| | | | | | | |
|----------------------------|------|---|-------------|--------|--------|-----------------|
| RESULT 7 | | AX114289 | 5952 bp | DNA | linear | PAT 11-MAY-2001 |
| LOCUS | | Sequence 1 from Patent WO0129243. | | | | |
| DEFINITION | | AX114289 | | | | |
| ACCESSION | | AX114289 | | | | |
| VERSION | | AX114289.1 | GI:14031259 | | | |
| KEYWORDS | | Homo sapiens (human) | | | | |
| SOURCE | | Homo sapiens | | | | |
| ORGANISM | | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | |
| REFERENCE | | 1 Paul,X.L. and Xiao,X. | | | | |
| AUTHORS | | Method and vector for producing and transferring trans-spliced peptides | | | | |
| TITLE | | Patent: WO 0129243-A 1 26-APR-2001; DALHOUSIE UNIVERSITY (CA) ; UNIV. OF PITTSBURGH OF THE COMMONWEALTH SYSTEM OF HIGHER EDUCATION (US) | | | | |
| JOURNAL | | Location/Qualifiers | | | | |
| FEATURES | | 1..5952 | | | | |
| source | | /organism="Homo sapiens" | | | | |
| | | /mol_type="genomic DNA" | | | | |
| | | /db_xref="taxon:9606" | | | | |
| misc_feature | | 2897..2898 | | | | |
| | | /note="S4 junction site" | | | | |
| misc_feature | | 3198..3199 | | | | |
| | | /note="S2 junction site" | | | | |
| BASE COUNT | | 1860 a | 1344 c | 1410 g | 1338 t | |
| ORIGIN | | | | | | |
| Query Match | | 50.1%; Score 1002.6; DB 6; Length 5952; | | | | |
| Best Local Similarity | | 99.6%; Pred. No. 7.8e-241; | | | | |
| Matches 1005; Conservative | | 0; Mismatches 4; Indels 0; Gaps 0; | | | | |
| QY | 993 | ACAGCACAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTC | 1052 | | | |
| Db | 3679 | ATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTC | 3738 | | | |
| QY | 1053 | AAGCTGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTT | 1112 | | | |
| Db | 3739 | AAGCTGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTT | 3798 | | | |
| QY | 1113 | GACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCATCTCGAGGAGAAAATTGGCGCTCTG | 1172 | | | |
| Db | 3799 | GACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCATCTCGAGGAGAAAATTGGCGCTCTG | 3858 | | | |
| QY | 1173 | AAAGAGAACGTGAGCCACCGTCATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAATTCAG | 1232 | | | |
| Db | 3859 | AAAGAGAACGTGAGCCACCGTCATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAATTCAG | 3918 | | | |
| QY | 1233 | CTCTCACCGTATAACCTCAGCACCTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAG | 1292 | | | |
| Db | 3919 | CTCTCACCGTATAACCTCAGCACCTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGCAG | 3978 | | | |
| QY | 1293 | GTGGCCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCA | 1352 | | | |
| Db | 3979 | GTGGCCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGACCTTTGGTCCAGCA | 4038 | | | |
| QY | 1353 | TCTCAGCATTTCTTTCCACGTCCTGTCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAAC | 1412 | | | |
| Db | 4039 | TCTCAGCATTTCTTTCCACGTCCTGTCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAAC | 4098 | | | |
| QY | 1413 | AAAGTGCCCTACTATATCAACCCAGAGACTCAAAACAACTTGTGGGAGAGCCATCCCAAATG | 1472 | | | |
| Db | 4099 | AAAGTGCCCTACTATATCAACCCAGAGACTCAAAACAACTTGTGGGAGAGCCATCCCAAATG | 4158 | | | |
| QY | 1473 | ACAGAGCTCTACCAAGTCTTTAGCTGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACT | 1532 | | | |
| Db | 4159 | ACAGAGCTCTACCAAGTCTTTAGCTGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACT | 4218 | | | |
| QY | 1533 | GCCATGAAACTCCGAAGACTGCAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCT | 1592 | | | |

| | | | |
|----------------------------|------|--|------|
| Db | 4219 | GCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCT | 4278 |
| QY | 1593 | GCATGTGATGCTTGGACCAAGCACAAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTG | 1652 |
| Db | 4279 | GCATGTGATGCTTGGACCAAGCACAAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTG | 4338 |
| QY | 1653 | CAGATTATTAATTGTTTGACCACTATTTATGACCGCCTGGAGCAAGAGACACAACAATTG | 1712 |
| Db | 4339 | CAGATTATTAATTGTTTGACCACTATTTATGACCGCCTGGAGCAAGAGACACAACAATTG | 4398 |
| QY | 1713 | GTCAACGTCCTCTCTGCGTGGATATGCTCTGAACTGGCTGCTGAATGTTTATGATACG | 1772 |
| Db | 4399 | GTCAACGTCCTCTCTGCGTGGATATGCTCTGAACTGGCTGCTGAATGTTTATGATACG | 4458 |
| QY | 1773 | GGACGAACAGGAGGATCCGTGTCCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAA | 1832 |
| Db | 4459 | GGACGAACAGGAGGATCCGTGTCCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAA | 4518 |
| QY | 1833 | GCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTT | 1892 |
| Db | 4519 | GCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTT | 4578 |
| QY | 1893 | TGTGACCAAGCGAGGCTGGCCTCCTTCTGATGATTTCTATCCAAATTTCAAGACAGTTG | 1952 |
| Db | 4579 | TGTGACCAAGCGAGGCTGGCCTCCTTCTGATGATTTCTATCCAAATTTCAAGACAGTTG | 4638 |
| QY | 1953 | GGTGAAGTTGCATCCTTTGGGGCGAGTAACATTGAGCCAAAGTGTCCGGA | 2001 |
| Db | 4639 | GGTGAAGTTGCATCCTTTGGGGCGAGTAACATTGAGCCAAAGTGTCCGGA | 4687 |
| RESULT 8 | | | |
| AX538622 | | | |
| LOCUS | | | |
| DEFINITION | | | |
| ACCESSION | | | |
| VERSION | | | |
| KEYWORDS | | | |
| SOURCE | | | |
| ORGANISM | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| FEATURES | | | |
| source | | | |
| BASE COUNT | | | |
| ORIGIN | | | |
| Query Match | | | |
| Best Local Similarity | | | |
| Matches 1005; Conservative | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| Db | | | |
| QY | | | |
| | | | |

```

/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"
BASE COUNT 3707 a 2339 c 2502 g 2895 t
ORIGIN

Query Match 50.1%; Score 1002.6; DB 6; Length 11443;
Best Local Similarity 99.6%; Pred. No. 8e-241;
Matches 1005; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 993 ACAGCACAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGCCACGGATGAGCTGGACCTC 1052
Db |||
Db 6479 ATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGCCACGGATGAGCTGGACCTC 6538
QY 1053 AAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATT 1112
Db |||
Db 6539 AAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATT 6598
QY 1113 GACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGGCCTCTG 1172
Db |||
Db 6599 GACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGGCCTCTG 6658
QY 1173 AAAGAGAACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTGAG 1232
Db |||
Db 6659 AAAGAGAACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTGAG 6718
QY 1233 CTCTCAGCGTATAACCTCAGCACCTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGAG 1292
Db |||
Db 6719 CTCTCAGCGTATAACCTCAGCACCTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGAG 6778
QY 1293 GTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCA 1352
Db |||
Db 6779 GTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCA 6838
QY 1353 TCTCAGCACTTTCTTTCCACGCTGTCTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAC 1412
Db |||
Db 6839 TCTCAGCACTTTCTTTCCACGCTGTCTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAC 6898
QY 1413 AAAGTGGCCTACTATATCAACACGAGACTCAAAACAATTGCTGGGACCATCCCAAAATG 1472
Db |||
Db 6899 AAAGTGGCCTACTATATCAACACGAGACTCAAAACAATTGCTGGGACCATCCCAAAATG 6958
QY 1473 ACAGAGCTCTACCAAGTCTTTAGTGACCTGAATAATGTGAGATTCTCAGCTTATAGGACT 1532
Db |||
Db 6959 ACAGAGCTCTACCAAGTCTTTAGTGACCTGAATAATGTGAGATTCTCAGCTTATAGGACT 7018
QY 1533 GCCATGAACCTCCGAAGACTGCGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCT 1592
Db |||
Db 7019 GCCATGAACCTCCGAAGACTGCGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCT 7078
QY 1593 GCATGTGATGCTTGGACCGACCAACCTCAAGCAAAATGACCGCCATGGATATCCTG 1652
Db |||
Db 7079 GCATGTGATGCTTGGACCGACCAACCTCAAGCAAAATGACCGCCATGGATATCCTG 7138
QY 1653 CAGATTATTAATTGTTTGACCACTATTATGACCGCCTGGAGCAAGAGCACAAATTTG 1712
Db |||
Db 7139 CAGATTATTAATTGTTTGACCACTATTATGACCGCCTGGAGCAAGAGCACAAATTTG 7198
QY 1713 GTCAACGTCCTCTCTGCGTGGATATGTGTGAACTGGCTGCTGAATGTTTATGATACG 1772
Db |||
Db 7199 GTCAACGTCCTCTCTGCGTGGATATGTGTGAACTGGCTGCTGAATGTTTATGATACG 7258
QY 1773 GGACGAAACAGGGAGGATCCGTGTCTCTTTTAAACTGGCATCATTTCCCTGTGTAA 1832
Db |||
Db 7259 GGACGAAACAGGGAGGATCCGTGTCTCTTTTAAACTGGCATCATTTCCCTGTGTAA 7318
QY 1833 GCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTT 1892
Db |||
Db 7319 GCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTT 7378
QY 1893 TGTGACCAAGCGCAGGCTGGGCCCTCTCTGCAATGATTTCCAAATTTCCAAAGACAGTTG 1952
Db |||
Db 7379 TGTGACCAAGCGCAGGCTGGGCCCTCTCTGCAATGATTTCCAAATTTCCAAAGACAGTTG 7438
```

```

QY 1953 GGTGAAGTTGCATCCTTTGGGGCAGTAACATTGAGCCAAGTGTCCGGA 2001
Db |||
Db 7439 GGTGAAGTTGCATCCTTTGGGGCAGTAACATTGAGCCAAGTGTCCGGA 7487

RESULT 10
AX538627 LOCUS AX538627 12057 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 47 from Patent WO0229056.
ACCESSION AX538627
VERSION AX538627.1 GI:25271181
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 artificial sequences.
AUTHORS Chamberlain,J.S. and Harper,S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 47 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
location/Qualifiers
source
1..12057
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"
BASE COUNT 4020 a 2456 c 2790 g 2791 t
ORIGIN

Query Match 50.1%; Score 1002.6; DB 6; Length 12057;
Best Local Similarity 99.6%; Pred. No. 8e-241;
Matches 1005; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 993 ACAGCACAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTC 1052
Db |||
Db 8993 ATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTC 9052
QY 1053 AAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTTGGCAGCCCGTGGCGCATCTCCTCATT 1112
Db |||
Db 9053 AAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTTGGCAGCCCGTGGCGCATCTCCTCATT 9112
QY 1113 GACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCCCTCTG 1172
Db |||
Db 9113 GACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCCCTCTG 9172
QY 1173 AAAGAGAACCGTGAGCCACGTCATGACCTTGTCCAGGCTTACCACTTTGGGCATTGAG 1232
Db |||
Db 9173 AAAGAGAACCGTGAGCCACGTCATGACCTTGTCCAGGCTTACCACTTTGGGCATTGAG 9232
QY 1233 CTCTCACCGTATAACCTCAGACACTCTTGAAGACCTTGAAGACCTGAAACACCAAGATGGAAGCTTCTGCAG 1292
Db |||
Db 9233 CTCTCACCGTATAACCTCAGACACTCTTGAAGACCTTGAAGACCTGAAACACCAAGATGGAAGCTTCTGCAG 9292
QY 1293 GTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGGACTTTGGTCCAGCA 1352
Db |||
Db 9293 GTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGGACTTTGGTCCAGCA 9352
QY 1353 TCTCAGCACTTTCTTTCCACGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAC 1412
Db |||
Db 9353 TCTCAGCACTTTCTTTCCACGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAC 9412
QY 1413 AAAGTGGCCTACTATATCAACACGAGACTCAAAACAATTGCTGGGACCATCCCAAAATG 1472
Db |||
Db 9413 AAAGTGGCCTACTATATCAACACGAGACTCAAAACAATTGCTGGGACCATCCCAAAATG 9472
QY 1473 ACAGAGCTCTACCAAGTCTTTAGTGACCTGAATAATGTGAGATTCTCAGCTTATAGGACT 1532
Db |||
Db 9473 ACAGAGCTCTACCAAGTCTTTAGTGACCTGAATAATGTGAGATTCTCAGCTTATAGGACT 9532
QY 1533 GCCATGAACCTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTTGAGCCTGTGAGCT 1592
Db |||
Db 9533 GCCATGAACCTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTTGAGCCTGTGAGCT 9592
```

Qy 1593 GCATGTGATGCCTTGGACGAGCACAAACCTCAAGCAAAATGACCAGCCCATGGATATCCTG 1652
|||||
Db 9593 GCATGTGATGCCTTGGACGAGCACAAACCTCAAGCAAAATGACCAGCCCATGGATATCCTG 9652
|||||
Qy 1653 CAGATTATTAAATTGTTTGACCACTATTATTATGACCGCCTGGAGCAAGACACAAATTTG 1712
|||||
Db 9653 CAGATTATTAAATTGTTTGACCACTATTATTATGACCGCCTGGAGCAAGACACAAATTTG 9712
|||||
Qy 1713 GTCAACGTCCTCTCTCGCTGGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACG 1772
|||||
Db 9713 GTCAACGTCCTCTCTCGCTGGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACG 9772
|||||
Qy 1773 GGACGAACAGGGAGGATCCGTGTCCTGTCTCTTTTAAACTGGCATCAATTTCCCTGTGTAAA 1832
|||||
Db 9773 GGACGAACAGGGAGGATCCGTGTCCTGTCTCTTTTAAACTGGCATCAATTTCCCTGTGTAAA 9832
|||||
Qy 1833 GCACATTTGGAAGACAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTT 1892
|||||
Db 9833 GCACATTTGGAAGACAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTT 9892
|||||
Qy 1893 TGTGACGACGCGAGGCTGGGCCTCCTTCTGCATGATCTATCCAAATTCAGACAGTTG 1952
|||||
Db 9893 TGTGACGACGCGAGGCTGGGCCTCCTTCTGCATGATCTATCCAAATTCAGACAGTTG 9952
|||||
Qy 1953 GGTGAAGTTGCATCCTTTGGGGGCAGTAACATTCAGCCCAAGTGTCCGGA 2001
|||||
Db 9953 GGTGAAGTTGCATCCTTTGGGGGCAGTAACATTCAGCCCAAGTGTCCGGA 10001
|||||

RESULT 11
HSDMDR Human mRNA for dystrophin. 12446 bp mRNA linear PRI 12-SEP-1993
LOCUS X14298
DEFINITION Dmd gene; Duchenne muscular dystrophy; dystrophin.
ACCESSION X14298.1 GI:30845
VERSION 1
KEYWORDS Dmd gene; Duchenne muscular dystrophy; dystrophin.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 12446)
Rosenthal,A., Speer,A., Billwitz,H., Cross,G.S., Forrest,S.M. and
Davies,K.E.
Two human cDNA molecules coding for the Duchenne muscular dystrophy
(DMD) locus are highly homologous
Nucleic Acids Res. 17 (13), 5391 (1989)
89345106
2668885
2 (bases 1 to 12446)
Rosenthal,A.
Direct Submission
Submitted (09-FEB-1989) Rosenthal A., Akademie der Wissenschaften
der DDR, Zentralinstitut fuer Molekularbiologie, Robert-Roessle
Str.10, 1115 Berlin Buch, DDR
see also M18533 and M20250 for Dmd seqs.; discrepancies compared to
M18533 cDNA were located at x14298 pos. 496, 1772, 1965, 2449,
3687, 4229, 4504, 5075, 5332, 5630 and 7194.
Location/Qualifiers
1. .12446
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X chromosomal, Xp21."
/tissue_type="muscle"
/dev_stage="fetal and adult."
99. .11156
/note="dystrophin (AA 1 - 3685)"
/codon_start=1
/protein_id="CAA32479.1"
/db_xref="GI:30846"
/db_xref="SWISS-PROT:P11532"
/translation="MLWEEVEDCYEREDVQKFTTKVWNAQFSKFGKHQHTENLFSDL

QDGRRLDLLLEGLTGQKLPEKKGSTRVHALLNNVNKALRVLQNNNVDLVNI GSTDIVD
NHKLTGLIWNIIILHWQVKNVMKNI MAGLQPTNSEKILLSWVROSTRNYPQVNVINFT
TSWSGDALNALIHSRHPDLFDWNSVVCQSATSATQLEHAFNIQLEKILLDPEDV
DTTYPDKKSIIMYITSLFOVL PQQVSIEAIQEVEMLP RPBPVKTKKEHFLQHMHQHSYQ
QITVSLAQGERTSSPKPRFKSYAYTAAVTTSDTPRFPFSQHLEAPEDKSGSSSL
MESEVNLDRYQTAL EEVLSWLLSAEDTLOAQGEISNDVEVVKDQFTHHEGYMMDLTAH
QGRVGNILQSGSLIGTKLSEDETEVQEOMNLLNSRWECLRVASMEKQSNLHRVLM
DLQNKLELNDWLTKTEERTKMEBEPGLPDLEDLKRQVQKHVLOEDLEQEQVRVN
SLTHMVVVVDESSGDHATAALEEQKVLGDRWANI CRWTEDRWLLQDILLKQWRLTE
EQCLFSAWLSEKEDAVNKIHTTGFKDQNEMLSSLOKLAVLADLEKKKQSMGKLYSIK
QDLLSTLKNKSVTKTEAWLDFNARCWDLNVQKLEKSTAQISOAVTTTPSLTQTITVM
ETVTTVTREQLLVKHAQEELPPPPQPKRQITVDSEIRKRLDVIDITELHSWITRSEA
VLQSPPEFAI FRKEGNFSDLKEKVNAIEREKAERFKLODASRSGOALVEQM VNEGNA
DSIKQASEQLNSRWIEFCQLLSERLWLEYQNNIIAFYNQLQQLBQMTTAAENWLKIQ
PTTPSEPTAIKSQLKICKDEVNRLSGLQPQIERLKIQSIALKEKGQGMFLDADFVAF
TNHFQVFSVQVAREKELQTIIFDILPMPRYQETMSAIRTVVQSETKLSIPQLSVDY
EIMEQRLGELQALQSSLOEQSGLYVLTSTTVKEMSKAPSEISRKYSQSEFEEIEGRWK
KLSSQLVEHCQKLEEQMKLRKIQNHQITLKKWMAEVDVFLKEEWPALGDS EILKKQL
KQCRLLVSDIQTIPSLNSVNEGOKTKNEAPEFASRLETELKELNTQWDHMCQQVY
ARKEALKGLEKTVSLQKDLSEMEHWMVQAEEYLERDFEYKTPDELQKAFEEMKRAK
EEAQQKEAGVKLLTESVNSVIAQAPVQAQEALEKLELTITNTYQWLCTRNLNGCKTLE
EVAWCWHELLSYLEKANKWLNEVEFKLTNTENIPGAAEISEVLDLSLENLMRHSNDP
NQIRILAQTITDGGVMDLINEELETFNSRWRELHEEAVRRQKLLQEQSIQSAQETENS
LHLIQESLTFIDKQLAAYIADKVDAAMPQEAQIKQSDLSHEISLEEMKKHQGKEA
AORVLSQIDVAQKQLQDVSMKFLFOKPFANFEQRLQESKMLDVEKMHLPALETKSVE
QEVVQSQNLHCNVLYKSLSEVKSEVMVITGRQIQVKKQTEPKELDERVTALKHY
NELGAKVTIERKQOLEKCLKSRKMKEMNVLTWLAATDMELTKRS AVEGMPSNLDSE
VAWGKATQKIEIKQVHLKSI TEVGEALKTVLGKETLVEDKLSLNSNWI AVTSRAE
EWNLLLEYQKHMETFDQNVDHITKWI IQADTL LDESEKKKPOQKEDVLKRLKAE LND
IRPKVDS TRDQAANLMANHGDCRKLVEPQISELNHRFAAISHRIKTKKASIPLKELE
QFNSDIQKLEPLEAEIQQGVNLKEEDFNKDMNEDNEGTVKELLQRGDNLQQRITDER
KSEEEKIKQOLLQTKHNALKDLSORRKKALEISHOWYQVKQADDLKCLDDIEKKL
ASLPEPRDERKIKEIDRELQKKKEELNAVRROAEGLSHOWYQVKQADDLKCLDDIEKKL
IESKFAQFERLNFQAHTVREETMMVMTDMPLEISYVPSVTYLTETHTVSOALLVEQ
LLNAPDLCAKDPEDLFKQEEESLKNIKDSLQSSGRIDIHSHKKTAAALQSATPVERVKL
QEA LSQDLQWEKVMKWKDRQGRFDRSVEKRRRHYDIKIFNQWLTEAEQFLRKTQI
PENWEHAKYKWLKELQDIGIGQRTVVRTLNATGEEIIQSSKT DASILEQKLGSLNL
RWQEVCKQLSDRKKRLEEOKNILSEFQRD LNEFVWLWEEADNIASIPLEPGKEQOLKE
KLEQVKLLVEELPLRQGI LKQNETGPPVLVSAPISPEEQDKLENKLTQTNLQWIKVS
RALPEKQGEIEAQIKDLGQLEKKLEDLEEQNLHLLWLSPINQLEIYINQPNQEGPFD
VKETEIAVQAKQPDVEEILSKGQHYKEKPATQPVKRLDLSSEKWKAVNRLLELRA
KOPDLAPGLTTIGASPTQTVTLVTPVVTKETAI SKLEMPSSLMLEVPALADFNRAWT
ELTDWLSLLDQVIKSQRMVMVDLEINEMIIKQKATMQDLEQRRPQLEELITAAQNLK
NKTSNQEARTIITDRIERIQNWDEVQEHQNRQQLNEMLKDSTQWLEAKEEAEQVL
GOARAKLESWKEGPYTVDAIQKKITETKQAKDLRQWQTNVDVANDLAKLLRDYSAD
DTRKVHMITENINASWRSIHKRVSEAEAL EETHRLLQQFPDLLEKFLAWL TEAETTA
NVLODAFTRKERLLEDKSGVKELMQQDLQGEIEAHTDVVHNDENSOKILRSLEGST
DAVLLQRLDNMNFKWSLRKKSINIRSHLEASSDQWKRLHLSLOELLVWLQKDDDEL
SROAPIGGDFPAVQKQNDVHRAPFKRELKTKEPVIMSTLETVRIFLTEQPLEGLEKLYQ
EPRELPPEEAQNVTLLRKRQABEVNTEWEKLNLSADWOKKIDETLERLQELQEATD
ELDLKLRQAEVIKGSWQPVGDILLDSLDQHLKVKALRGEIAPLKENVSHVNDLARQL
TTLGIOQSPYNLSTLEDLNRWKLQVAVEDRVRQLHEAHRDFGPASQHFSTSVQGP
WERAISPNKVPPYINHETQTTCTWDHPKMTELYQSLADLNVRFFSAYRTAMKLRRLQKA
LCDLLLSAACDALDQHNLKONDQPM DILQIINCLTTIYDRLEQEHNNLVNPLCVD
MCLNWLNLVYDTRTGTRIRVLSFKTGIISLCKAHLDEKRYLFFKQVASSTGFCDDQRL
GLLHDSIQIPRQLGEVASFGGSNI EPSVRSCTCFANNKPEIEAALFLDWMRLPQSM
VWLPVLHRVAAAEATAKHAQKCNICECPIIGFRYRSLKHFNVDICQSCFFSGRVAKGH
KMHPMVEYCTPTTSGEDVRDFAKVLKNKFRTRKYFAKHPRMGYLPVQTVLEGDNMET
PVTLLNFWPVD SAPASSPQLSHDDTHSRIEHYASRLAEMNSNGSYLND SISPNESID
DEHLIQHYCQSLNQDSPLSQRSPAQILISLESEERGELEIRILADLEENRNLOAEY
DRLKQHEHKGLSPLSPPEMPTSPQSPRDAELIAEAKLLRQHKGRLEARMQILEDH
NKQESQLHRLRQLLEQPOAEAKVNGTTVSSPSTSLQRDSSQPMLLRLLRVVGSQTSDSM
GEEDLLSPPDQDSTSTGLEEVMBQLNNSFPSSRGRNTPGKPNREDTM"

BASE COUNT 4135 a 2524 c 2876 g 2911 t
ORIGIN

Query Match 50.1%; Score 1002.6; DB 9; Length 12446;
Best Local Similarity 99.6%; Pred. No. 8e-241;
Matches 1005; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 993 ACAGCACAGACCCCTTGAAGAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTC 1052
|||
Db 8893 ATAGATGAGACCCCTTGAAGAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTC 8942

QY 1053 AAGCTGCGCCCAAGTGAGGTGATCAAGGGATCCTGGCAGCCCCGTGGCGGATCTCCTCATTT 1112
Db |||||
8943 AAGCTGCGCCCAAGTGAGGTGATCAAGGGATCCTGGCAGCCCCGTGGCGGATCTCCTCATTT 9002
QY 1113 GACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGAGAGAAATTGGCGCTCTG 1172
Db |||||
9003 GACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGAGAGAAATTGGCGCTCTG 9062
QY 1173 AAAGAGAACGTCAGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAG 1232
Db |||||
9063 AAAGAGAACGTCAGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAG 9122
QY 1233 CTCTCACCGTATAACCTCAGCACCTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGAG 1292
Db |||||
9123 CTCTCACCGTATAACCTCAGCACCTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGAG 9182
QY 1293 GTGGCCGTCGAGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCA 1352
Db |||||
9183 GTGGCCGTCGAGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCA 9242
QY 1353 TCTCAGCACCTTTCTTTCCACGTCGTCTCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAAC 1412
Db |||||
9243 TCTCAGCACCTTTCTTTCCACGTCGTCTCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAAC 9302
QY 1413 AAAGTGCCCTACTATATCAACACCGAGACTCAAAACAACTTGCTGGGACCATCCCAAATG 1472
Db |||||
9303 AAAGTGCCCTACTATATCAACACCGAGACTCAAAACAACTTGCTGGGACCATCCCAAATG 9362
QY 1473 ACAGAGCTCTACAGTCTTTAGCTGACCTGAATAATGTGAGATCTCAGCTTATAGGACT 1532
Db |||||
9363 ACAGAGCTCTACAGTCTTTAGCTGACCTGAATAATGTGAGATCTCAGCTTATAGGACT 9422
QY 1533 GCCATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTGAGCCTGTGAGCT 1592
Db |||||
9423 GCCATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTGAGCCTGTGAGCT 9482
QY 1593 GCATGTGATGCCCTGGACCAGCACAACTCAAGCAAAATGACAGCCCATGGATATCCTG 1652
Db |||||
9483 GCATGTGATGCCCTGGACCAGCACAACTCAAGCAAAATGACAGCCCATGGATATCCTG 9542
QY 1653 CAGATTATTAATTGTTGACCACCTATTTATGACCGCCTGGAGCAAGAGACACAACAATTTG 1712
Db |||||
9543 CAGATTATTAATTGTTGACCACCTATTTATGACCGCCTGGAGCAAGAGACACAACAATTTG 9602
QY 1713 GTCAACGTCCTCTCTGCGTGGATATGTGTGAACTGGCTGCTGAATGTTTATGATACG 1772
Db |||||
9603 GTCAACGTCCTCTCTGCGTGGATATGTGTGAACTGGCTGCTGAATGTTTATGATACG 9662
QY 1773 GGACGAACAGGGAGGATCCGTGTCTCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAA 1832
Db |||||
9663 GGACGAACAGGGAGGATCCGTGTCTCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAA 9722
QY 1833 GCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTT 1892
Db |||||
9723 GCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTT 9782
QY 1893 TGTGACCAAGCGAGCTGGGCCCTCCTTCTGCATGATTTCTATCCAAATTTCAAGACAGTTG 1952
Db |||||
9783 TGTGACCAAGCGAGCTGGGCCCTCCTTCTGCATGATTTCTATCCAAATTTCAAGACAGTTG 9842
QY 1953 GGTGAAGTTGCATCCTTTGGGGCAGTAACATTTAGCCCAAGTGTCCGGA 2001
Db |||||
9843 GGTGAAGTTGCATCCTTTGGGGCAGTAACATTTAGCCCAAGTGTCCGGA 9891

RESULT 12
AX409637
LOCUS AX409637 13957 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 2284 from Patent WO229103.
ACCESSION AX409637
VERSION AX409637.1 GI:21442342
KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2284 11-APR-2002;
GENE LOGIC INC (US)
FEATURES Location/Qualifiers
source
1. .13957
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. M18533"
BASE COUNT 4602 a 2781 c 3122 g 3452 t
ORIGIN
Query Match 50.1%; Score 1002.6; DB 6; Length 13957;
Best Local Similarity 99.6%; Pred. No. 8e-241;
Matches 1005; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 993 ACAGCACAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTC 1052
Db |||||
8993 ATAGATGAGACCCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTC 9052
QY 1053 AAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTT 1112
Db |||||
9053 AAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTT 9112
QY 1113 GACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGAGAGAAATTGGCGCTCTG 1172
Db |||||
9113 GACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGAGAGAAATTGGCGCTCTG 9172
QY 1173 AAAGAGAACGTCAGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAG 1232
Db |||||
9173 AAAGAGAACGTCAGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAG 9232
QY 1233 CTCTCACCGTATAACCTCAGCACCTTCGAAAGACCTGAACACCCAGATGGAAGCTTCTGCAG 1292
Db |||||
9233 CTCTCACCGTATAACCTCAGCACCTTCGAAAGACCTGAACACCCAGATGGAAGCTTCTGCAG 9292
QY 1293 GTGGCCGTCGAGGACCGAGTCAGGACGCTGCATGAAGCCCAAGGCCACCTTCGCCAAAC 1352
Db |||||
9293 GTGGCCGTCGAGGACCGAGTCAGGACGCTGCATGAAGCCCAAGGCCACCTTCGCCAAAC 9352
QY 1353 TCTCAGCACCTTTCTTTCCACGTCGTCTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAAC 1412
Db |||||
9353 TCTCAGCACCTTTCTTTCCACGTCGTCTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAAC 9412
QY 1413 AAAGTGCCCTACTATATCAACCCAGAGACTCAAAACAACTTGCTGGGACCATCCCAAATG 1472
Db |||||
9413 AAAGTGCCCTACTATATCAACCCAGAGACTCAAAACAACTTGCTGGGACCATCCCAAATG 9472
QY 1473 ACAGAGCTCTACCAAGTCTTTAGCTGACCTGAATAATGTGAGATCTCAGCTTATAGGACT 1532
Db |||||
9473 ACAGAGCTCTACCAAGTCTTTAGCTGACCTGAATAATGTGAGATCTCAGCTTATAGGACT 9532
QY 1533 GCCATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTTTGAGCCTGTGAGCT 1592
Db |||||
9533 GCCATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTTTGAGCCTGTGAGCT 9592
QY 1593 GCATGTGATGCTTGGACCAGCACAACTCAAGCAAAATGACAGCCCATGGATATCCTG 1652
Db |||||
9593 GCATGTGATGCTTGGACCAGCACAACTCAAGCAAAATGACAGCCCATGGATATCCTG 9652
QY 1653 CAGATTATTAATTGTTGACCACCTATTTATGACCGCCTGGAGCAAGAGACACAACAATTTG 1712
Db |||||
9653 CAGATTATTAATTGTTGACCACCTATTTATGACCGCCTGGAGCAAGAGACACAACAATTTG 9712
QY 1713 GTCAACGTCCTCTCTGCGTGGATATGTGTGAACTGGCTGCTGAATGTTTATGATACG 1772
Db |||||
9713 GTCAACGTCCTCTCTGCGTGGATATGTGTGAACTGGCTGCTGAATGTTTATGATACG 9772

QY 1773 GGACGAACAGGAGGATCCGTGTCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAA 1832
Db 9773 GGACGAACAGGAGGATCCGTGTCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAA 9832
QY 1833 GCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTT 1892
Db 9833 GCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTT 9892
QY 1893 TGTGACCAGCGAGCTGGGCCTCCTTCTGCGATGATTCTATCCAAATTCCAAGACAGTTG 1952
Db 9893 TGTGACCAGCGAGCTGGGCCTCCTTCTGCGATGATTCTATCCAAATTCCAAGACAGTTG 9952
QY 1953 GGTGAAGTTGCATCCTTTTGGGGGCAGTAACATTGAGCCAAAGTGTCCGGA 2001
Db 9953 GGTGAAGTTGCATCCTTTTGGGGGCAGTAACATTGAGCCAAAGTGTCCGGA 10001

RESULT 13
AX538581
LOCUS AX538581 13957 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 1 from Patent WO0229056.
ACCESSION AX538581
VERSION AX538581.1 GI:25271086
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Chamberlain,J.S. and Harper,S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 1 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
source
1. .13957
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 4602 a 2781 c 3122 g 3452 t
ORIGIN

Query Match 50.1%; Score 1002.6; DB 6; Length 13957;
Best Local Similarity 99.6%; Pred. No. 8e-241;
Matches 1005; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 993 ACAGCACAGACCCCTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTC 1052
Db 8993 ATAGATGAGACCCCTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTC 9052
QY 1053 AAGCTGCGCCAAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTT 1112
Db 9053 AAGCTGCGCCAAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTT 9112
QY 1113 GACTCTCTCCAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTG 1172
Db 9113 GACTCTCTCCAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTG 9172
QY 1173 AAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTGAG 1232
Db 9173 AAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTGAG 9232
QY 1233 CTCTCACCCTATACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAG 1292
Db 9233 CTCTCACCCTATACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAG 9292
QY 1293 GTGGCCGTGAGGACCGAGTCAGGCAGCTGATGAAGCCCAAGGGACCTTTGGTCCAGCA 1352
Db 9293 GTGGCCGTGAGGACCGAGTCAGGCAGCTGATGAAGCCCAAGGGACCTTTGGTCCAGCA 9352
QY 1353 TCTCAGCACTTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGGCCATCTCGCCAAAC 1412
Db 9353 TCTCAGCACTTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGGCCATCTCGCCAAAC 9412

QY 1413 AAAGTCCCCTACTATATCAACCACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAATG 1472
Db 9413 AAAGTCCCCTACTATATCAACCACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAATG 9472
QY 1473 ACAGAGCTCTACCACTCTTTTAGCTGACCTGAAATAATGTGAGATTCTCAGCTTATAGGACT 1532
Db 9473 ACAGAGCTCTACCACTCTTTTAGCTGACCTGAAATAATGTGAGATTCTCAGCTTATAGGACT 9532
QY 1533 GCCATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCT 1592
Db 9533 GCCATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCT 9592
QY 1593 GCATGTGATGCTTGGACCAAGCAACACCTCAAGCAAAATGACCAAGCCCATGGATATCCTG 1652
Db 9593 GCATGTGATGCTTGGACCAAGCAACACCTCAAGCAAAATGACCAAGCCCATGGATATCCTG 9652
QY 1653 CAGATTATTAAATTGTTGACCACTATTATTGACCCGCTGGAGCAAGAGCAACAATTTG 1712
Db 9653 CAGATTATTAAATTGTTGACCACTATTATTGACCCGCTGGAGCAAGAGCAACAATTTG 9712
QY 1713 GTCAACGTCCCTCTCTGCGTGGATATGTGCTGAACTGGCTGCTGAATGTTTATGATACG 1772
Db 9713 GTCAACGTCCCTCTCTGCGTGGATATGTGCTGAACTGGCTGCTGAATGTTTATGATACG 9772
QY 1773 GGACGAACAGGAGGATCCGTGTCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAA 1832
Db 9773 GGACGAACAGGAGGATCCGTGTCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAA 9832
QY 1833 GCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTT 1892
Db 9833 GCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTT 9892
QY 1893 TGTGACCAGCGAGGCTGGGCCTCCTTCTGCGATGATTCTATCCAAATTCCAAGACAGTTG 1952
Db 9893 TGTGACCAGCGAGGCTGGGCCTCCTTCTGCGATGATTCTATCCAAATTCCAAGACAGTTG 9952
QY 1953 GGTGAAGTTGCATCCTTTTGGGGGCAGTAACATTGAGCCAAAGTGTCCGGA 2001
Db 9953 GGTGAAGTTGCATCCTTTTGGGGGCAGTAACATTGAGCCAAAGTGTCCGGA 10001

RESULT 14
HUMDYS
LOCUS HUMDYS 13957 bp mRNA linear PRI 25-MAY-2000
DEFINITION Homo sapiens dystrophin (DMD) mRNA, complete cds.
ACCESSION M18533 M17154 M18026 M20250
VERSION M18533.1 GI:181856
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Koenig,M., Hoffman,E.P., Bertelson,C.J., Monaco,A.P., Feener,C. and Kunkel,L.M.
TITLE Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals
JOURNAL Cell 50 (3), 509-517 (1987)
MEDLINE 87273512
PUBMED 3607877
REFERENCE 2
AUTHORS Hoffman,E.P., Monaco,A.P., Feener,C.C. and Kunkel,L.M.
TITLE Conservation of the Duchenne muscular dystrophy gene in mice and humans
JOURNAL Science 238 (4825), 347-350 (1987)
MEDLINE 88018015
PUBMED 3659917
REFERENCE 3
AUTHORS Koenig,M., Monaco,A.P. and Kunkel,L.M.
TITLE The complete sequence of dystrophin predicts a rod-shaped cytoskeletal protein

JOURNAL
MEDLINE
PUBMED
COMMENT

Cell 53 (2), 219-226 (1988)
88194521
3282674

On May 25, 2000 this sequence version replaced gi:340693.
Draft entry and computer-readable sequence kindly provided by
M.Koenig, 01-APR-1988 The severity of muscular dystrophy is
determined by the size of the deleted DNA segment. Deletions found
in different patients were from positions 302-2200, 473-1168,
1691-1810, and 1169-3011.

FEATURES

Location/Qualifiers
1. .13957
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp21.3-p21.1"
/tissue_type="muscle"
/dev_stage="fetus"
/note="G00-119-850"
1. .13957
/gene="DMD"
209. .11266
/gene="DMD"
/codon_start=1
/product="dystrophin"
/protein_id="AAA53189.1"
/db_xref="GI:181857"
/translation="MLWVEEVEDCYEREDVQKKTFTKWVNAQFSKFGKOHENLFSDL
QDGRLLDLLEGLTGQKLPEKGGSTRVHALNNVNKALRVLQNNVDLVNIGSTDIVDG
NHKLTGLIWNIIHLWQVKNVKNIMAGLQQTSEKILLSWVRQSTRNYPQVNVINFT
TSWSDGLALNALIHSRPLDFDWSVVCQSATORLEHAFNIARYQLGIEKLLDPEDV
DITVPDKSILMYITSLFOVLPQVSIIEAIOEVEMLPRPPKVTKEEHFQLHHOMHSQ
QITVSLAQGYERTSSPKPRFKSYAYTAAAYVTTSDPFRSPPPSQHLEAPEDKSGSSSL
MESEVNLDRYQTALREVLSWLLSAEDTLQAQGEISNDVEVVKDQFHTHEGYMMDLTAH
QGRVGNILQGSKLIGTGKLSDEDETEVQEQMNLNSRWECLRVASMEKOSNLHRVLM
DLNQKLIKELNDWLTKTEERTPRKMEEPPLGPDLELKRQVQKHVLQEDLEQEQVRVN
SLTHMVVVDESSEGDHATAALEEQLKVLGDRWANI CRWTEDRWVLLQDILLKWRQLTE
EQCLFSAWLSKEKEDAVNKIHTTGFKDQNEMLSSLQKLAVLKADLEKKKQSGNKLKYLK
QDLSLTNKNKSVTQKTEAWLDPNFCWDLNVQKLEKSTAQISQAVTTTPSLTQTTVM
ETVTTVTTREQILVHAQEELPPPPPKKQROIITVDSIRKRLDVIDITELHSWITRSEA
VLSQPEFAIFRKEGNFSDLKEKNNAIEREKAERFKRLQDASRSQAALVEQVMNEGVNA
DSIQASEQLNSRWIEFCQLLSERLNWLEYQNNIIAFNQLQQLQLEQMTTAEWLKIQ
PTTPSEPTAIKSQLKICKDEVNRLSGLQPQIERLKIQSIALKEKGQGMFLDADFVAF
TNHFQVFSVDQAREKELQTIPTLPPMRYQETMSAIRTWVQOSETKLSIPQLSVTDY
EIMEORLQELQALQSSLOEQSGLYLSTTVKMSKAPSEISRKYQSEFEIEGRWK
KLSSQLVEHQKLEEQMNLKRIQNHIOQLKKMAEVDVFLKEEWPALNGDSEILLKKQL
KQRLIVSDIQTIQPSLNSVNEGQIKNEAEPEFAGRLTELKELNLTQNDHMCQQVY
ARKEALGGLEKTVSLQKLSMHEWMTQAESEYLERDFEYKTPDELQKVAEEMKRAK
EEAQCKEAKVKLLTESVNSVIAQAPPVAQEAALKKELETLTNYOWLCTRLNGCKCTLE
EVNACWHELKLSYLEKANKWLNVEVEFKLKTENIPGGABEISEVLDSLENMRHSEDNP
NOIRILACTLTLDGVMDELINBELETFNSRWRELHEBARVRKQLLEQSIQASETEKS
LHLIQESLTFIDKQLAAYIADKVDAAMPQEAQKIQSDLTSHSISLEEMKKHQKKEA
AQVLSQIDVAQKKIQDVSMMKFLFQKPAFELRLQSKMILDEVKMHLPALETKSVE
QEVVQSLNHCNVNLYKSLSEVSEVEMVIKTRQIVQKKQNTENPKELDERVTLKLVHY
NELGAKVTERKQQLKCLKLSRKMRKEMNVLTWELAATDMELTKRSAVEGMPNSLDSE
VAWGKATQKEIEKQKHLKSTIEVGEALKTVLGKKEITLVEDKLSLLNSNWTAVTSRAE
EWNLLLELYQKHMETFDQNVDHITKWI IQADTLLDESEKKKPPQOKEDVLKRLKAEIND
IRPKVDSFTRDQAANLNRGDHCRKLVEPQISELNDHFPAAISHRIKTKGASTPLKELE
QFNSDIQKLLPLEAEIQQGVNLKEEDFNKDMNEDNEGTVKELLQRGDNLQQRITDER
KREEIKTKQQLQTKHNALKDLFSQRRKKALEISHQWYQYKRDADLLKCLDDIEKKL
ASLPEPRDERKIKEIDRELQKKEELNAVRRQAEGLSEDGAAMAVEPTQIQLSKRWE
IESKFAQPRRLNFAQIHTVREETNMVMTEDMPLEISYVPSYLTLEITHVSQALLEVEQ
LLNAPDLCAKDFEDLFKQEEESLNKIDSLQOSSGRIDIHSHKTAALQASATPVERVKL
QEALSQLDFQWEKVNKMYKDRGFRDRSVEKWRFRHYDIKIFNQLWTEAEQFLRKTQI
PENWEHAKYKWLKELQDGIQIGORQTVVRLNATGEEIIQQSSKTDASILQELGSLNL
RWQEVCKQLSDRKRLKEEQKNILSEFQDLNEFVLWLEADNIIASIPLEPGKBOQLKE
KLEQVKLVEELPLRQGIKQLNETGGPVLVSAPISPEEQDKLENKLNQTNLQWIKVS
RALPEKQGEIEAQIKDLGOLEKLELDEEQLNHLIISPIRNOLEIYNQPNQBGPPFD
VQETEIAVQAKQPDVEEILSKGHLYKEKPATQPVKRLLEDLSSEKAVNRLQLQELRA
KQPDAPGLTTIGASPTQTVTLVTPQVVTKETAI SKLEMPSSLMLEVPALADFNRAWT
ELTDWLSLLDQVIKSORVMVGLEDINEMI IKQKATMQDLEQRRPQLEELTAAQNKLK
NKTSNQEARTIITDIERIQNWDEVQEHQNRQQLNEMLKDSTQWLEAKEAEQVL
GQARAKLESWKEGPYTVDAIQKKITETKQAKDLRQWQTNVVDVANDLALKLRDYSAD

gene
CDS

| | | | | | |
|-----------------------|------|--|-------------------|------------|---------------|
| Query Match | | 50.1%; | Score 1002.6; | DB 9; | Length 13957; |
| Best Local Similarity | | 99.6%; | Pred. No. 8e-241; | | |
| Matches 1005; | | Conservative | 0; | Mismatches | 4; |
| | | | | Indels | 0; |
| | | | | Gaps | 0; |
| QY | 993 | ACAGCACAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTC | 1052 | | |
| Db | 8993 | ATAGATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTC | 9052 | | |
| QY | 1053 | AAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGCATCTCCTCATT | 1112 | | |
| Db | 9053 | AAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGCATCTCCTCATT | 9112 | | |
| QY | 1113 | GACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAAATTGCGCCTCTG | 1172 | | |
| Db | 9113 | GACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAAATTGCGCCTCTG | 9172 | | |
| QY | 1173 | AAAGAGAACGTGAGCCACGTCATGACCTTGTGCGCAGCTTACCACTTTGGGCATTTCAG | 1232 | | |
| Db | 9173 | AAAGAGAACGTGAGCCACGTCATGACCTTGTGCGCAGCTTACCACTTTGGGCATTTCAG | 9232 | | |
| QY | 1233 | CTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACACGATGGAAGCTTCTGCAG | 1292 | | |
| Db | 9233 | CTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACACGATGGAAGCTTCTGCAG | 9292 | | |
| QY | 1293 | GTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCA | 1352 | | |
| Db | 9293 | GTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCA | 9352 | | |
| QY | 1353 | TCTCAGCACTTTCTTTCCAGCTCTGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAAC | 1412 | | |
| Db | 9353 | TCTCAGCACTTTCTTTCCAGCTCTGTCCAGGGTCCCTGGGAGAGAGCCCATCTCGCCAAAC | 9412 | | |
| QY | 1413 | AAAGTGCCCTACTATATCAACACGAGACTCAAAACAACTTGTGTGGAGCCATCCCAAAATG | 1472 | | |
| Db | 9413 | AAAGTGCCCTACTATATCAACACGAGACTCAAAACAACTTGTGTGGAGCCATCCCAAAATG | 9472 | | |
| QY | 1473 | ACAGAGCTCTACCACTCTTTAGCTGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACT | 1532 | | |
| Db | 9473 | ACAGAGCTCTACCACTCTTTAGCTGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACT | 9532 | | |
| QY | 1533 | GCCATGAAACTCCGAAGACTGCAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAAGCT | 1592 | | |
| Db | 9533 | GCCATGAAACTCCGAAGACTGCAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAAGCT | 9592 | | |
| QY | 1593 | GCATGTGATGCCCTTGGACCAGCACAACTCAAGCAAAAATGACAGCCCATGGATATCCTG | 1652 | | |
| Db | 9593 | GCATGTGATGCCCTTGGACCAGCACAACTCAAGCAAAAATGACAGCCCATGGATATCCTG | 9652 | | |
| QY | 1653 | CAGATTATTAAATGTTGACCACCTATTATGACCGCCTGGAGCAAGAGCACAAATTTG | 1712 | | |
| Db | 9653 | CAGATTATTAAATGTTGACCACCTATTATGACCGCCTGGAGCAAGAGCACAAATTTG | 9712 | | |

| | | | |
|---|------|--|-------|
| QY | 1713 | GTCAACGTCCTCTCTGCGTGGATATGTGTCGAACTGGCTGCTGAATGTTTATGATACG | 1772 |
| Db | 9713 | GTCAACGTCCTCTCTGCGTGGATATGTGTCGAACTGGCTGCTGAATGTTTATGATACG | 9772 |
| QY | 1773 | GGACGAACAGGGAGGATCCGTGTCTCTGCTTTTAAAACTGGCATCATTTCCCTGTGTAAA | 1832 |
| Db | 9773 | GGACGAACAGGGAGGATCCGTGTCTCTGCTTTTAAAACTGGCATCATTTCCCTGTGTAAA | 9832 |
| QY | 1833 | GCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTT | 1892 |
| Db | 9833 | GCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTT | 9892 |
| QY | 1893 | TGTGACCAGCGCAGGTGGCCCTCCTTCTGTCATGATTTCTATCCAAATTTCCAAGACAGTTG | 1952 |
| Db | 9893 | TGTGACCAGCGCAGGTGGCCCTCCTTCTGTCATGATTTCTATCCAAATTTCCAAGACAGTTG | 9952 |
| QY | 1953 | GGTGAAGTTGCATCCTTTTGGGGGCAGTAACATTTAGCCCAAGTGTCCGGA | 2001 |
| Db | 9953 | GGTGAAGTTGCATCCTTTTGGGGGCAGTAACATTTAGCCCAAGTGTCCGGA | 10001 |
| RESULT 15 | | | |
| AR220819 | | | |
| LOCUS | | | |
| DEFINITION | | | |
| Sequence 60 from patent US 6426186. | | | |
| AR220819 | | | |
| ACCESSION | | | |
| AR220819.1 | | | |
| VERSION | | | |
| GI:23327696 | | | |
| KEYWORDS | | | |
| SOURCE | | | |
| Unknown. | | | |
| ORGANISM | | | |
| Unknown. | | | |
| REFERENCE | | | |
| 1 (bases 1 to 13977) | | | |
| Unclassified. | | | |
| Jones, K.A., Volkmut, W. and Walker, M.G. | | | |
| Bone remodeling genes | | | |
| Patent: US 6426186-A 60 30-JUL-2002; | | | |
| JOURNAL | | | |
| Location/Qualifiers | | | |
| source | | | |
| 1. 13977 | | | |
| /organism="unknown" | | | |
| BASE COUNT | | | |
| 4596 a 2765 c 3120 g 3453 t 43 others | | | |
| ORIGIN | | | |
| Query Match 49.6%; Score 991.6; DB 6; Length 13977; | | | |
| Best Local Similarity 99.5%; Pred. No. 4.7e-238; | | | |
| Matches 1005; Conservative 0; Mismatches 4; Indels 1; Gaps 1; | | | |
| QY | 993 | ACAGCACAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCGACGGATGAGCTGGACCTC | 1052 |
| Db | 8993 | ATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCGACGGATGAGCTGGACCTC | 9052 |
| QY | 1053 | AAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATT | 1112 |
| Db | 9053 | AAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATT | 9112 |
| QY | 1113 | GACTCTCTCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCTCTG | 1172 |
| Db | 9113 | GACTCTCTCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCTCTG | 9172 |
| QY | 1173 | AAAGAGAACCTGAGCCACGTCATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAATTCAG | 1232 |
| Db | 9173 | AAAGAGAACCTGAGCCACGTCATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAATTCAG | 9232 |
| QY | 1233 | CTCTCACCGTATAACCTCAGCACTCTGGAAGACCTTGACACCAGATGGAAGCTTCTGCAG | 1292 |
| Db | 9233 | CTCTCACCGTATAACCTCAGCACTCTGGAAGACCTTGACACCAGATGGAAGCTTCTGCAG | 9292 |
| QY | 1293 | GTGGCCGTCGAGGACCGAGTCAGGCGAGCTGCATGAAGCCCAAGGGACTTTGGTCCAGCA | 1352 |
| Db | 9293 | GTGGCCGTCGAGGACCGAGTCAGGCGAGCTGCATGAAGCCCAAGGGACTTTGGTCCAGCA | 9352 |
| QY | 1353 | TCTCAGCACTTTCTTTCCACGTCGTCTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAC | 1412 |
| Db | 9353 | TCTCAGCACTTTCTTTCCACGTCGTCTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAC | 9412 |

| | | | |
|----|------|--|-------|
| QY | 1413 | AAAGTGCCCTACTATATCAACACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAATG | 1472 |
| Db | 9413 | AAAGTGCCCTACTATATCAACACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAATG | 9472 |
| QY | 1473 | ACAGAGCTCTACCACTCTTTAGCTGACCTGAATAATGTAGATTCTCAGCTTATAGGACT | 1532 |
| Db | 9473 | ACAGAGCTCTACCACTCTTTAGCTGACCTGAATAATGTAGATTCTCAGCTTATAGGACT | 9532 |
| QY | 1533 | GCCATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTAGCT | 1592 |
| Db | 9533 | GCCATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTAGCT | 9592 |
| QY | 1593 | GCATGTGATGCTTGGACCAAGCAACACCTCAAGCAAAATGACCAAGCCATGGATATCCTG | 1652 |
| Db | 9593 | GCATGTGATGCTTGGACCAAGCAACACCTCAAGCAAAATGACCAAGCCATGGATATCCTG | 9652 |
| QY | 1653 | CAGATTATTAAATGTTTGACCACTATTATGACCGCCTGGAGCAAGAGCACAACAATTG | 1712 |
| Db | 9653 | CAGATTATTAAATGTTTGACCACTATTATGACCGCCTGGAGCAAGAGCACAACAATTG | 9712 |
| QY | 1713 | GTCAACGTCCTCTCTGCGTGGATATGTGTCGAACTGGCTGCTGAATGTTTATGATACG | 1772 |
| Db | 9713 | GTCAACGTCCTCTCTGCGTGGATATGTGTCGAACTGGCTGCTGAATGTTTATGATACG | 9772 |
| QY | 1773 | GGACGAACAGGGAGGATCCGTGTCCTGCTTTTAAAACTGGCATCATTTCCCTGTGTAAA | 1832 |
| Db | 9773 | GGACGAACAGGGAGGATCCGTGTCCTGCTTTTAAAACTGGCATCATTTCCCTGTGTAAA | 9832 |
| QY | 1833 | GCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTT | 1892 |
| Db | 9833 | GCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTT | 9892 |
| QY | 1893 | TGTGACCAGCGCAGGCT-GGGCCTCCTTCTGTCATGATTCTATCCAAATTTCCAAGACAGTT | 1951 |
| Db | 9893 | TGTGACCAGCGCAGGCTGGGGCCTCCTTCTGTCATGATTCTATCCAAATTTCCAAGACAGTT | 9952 |
| QY | 1952 | GGGTGAAGTTGCATCCTTTGGGGGCAGTAACATTGAGCCCAAGTGTCCGGA | 2001 |
| Db | 9953 | GGGTGAAGTTGCATCCTTTGGGGGCAGTAACATTGAGCCCAAGTGTCCGGA | 10002 |

Search completed: February 1, 2004, 23:51:35
Job time : 7142.37 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 10:13:25 ; Search time 475.292 Seconds
(without alignments)
11364.749 Million cell updates/sec

Title: US-09-845-416-14_COPY_1000_3000
Perfect score: 2001
Sequence: 1 ggcagttcattgatggagag.....cattgagccaagtgtccgga 2001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|--------------------|
| 1 | 2001 | 100.0 | 3446 | 24 | Human dystrophin m |
| 2 | 1990 | 99.5 | 4414 | 24 | Adeno-associated v |
| 3 | 1820.8 | 91.0 | 5417 | 24 | DNA encoding mini- |
| 4 | 1593 | 79.6 | 3510 | 24 | Human dystrophin m |
| 5 | 1593 | 79.6 | 4476 | 24 | Adeno-associated v |
| 6 | 1332 | 66.6 | 4402 | 21 | A rod shortened dy |
| 7 | 1269.2 | 63.4 | 4402 | 21 | A rod shortened dy |
| 8 | 1263 | 63.1 | 3858 | 24 | Human dystrophin m |

| | | | | | | |
|----|--------|------|-------|----|----------|--------------------|
| 9 | 1263 | 63.1 | 4825 | 24 | AAD37257 | Adeno-associated v |
| 10 | 1263 | 63.1 | 4848 | 24 | AAD37263 | Adeno-associated v |
| 11 | 1263 | 63.1 | 5060 | 24 | AAD37264 | Adeno-associated v |
| 12 | 1254 | 62.7 | 5339 | 24 | ABK81998 | DNA encoding mini- |
| 13 | 1251.2 | 62.5 | 3531 | 24 | AAD37238 | Human dystrophin m |
| 14 | 1251.2 | 62.5 | 4498 | 24 | AAD37258 | Adeno-associated v |
| 15 | 1245 | 62.2 | 4182 | 24 | AAD37230 | Human dystrophin m |
| 16 | 1245 | 62.2 | 5149 | 24 | AAD37255 | Adeno-associated v |
| 17 | 1132.2 | 56.6 | 5462 | 24 | ABK81999 | DNA encoding mini- |
| 18 | 1112 | 55.6 | 3999 | 24 | AAD37234 | Human dystrophin m |
| 19 | 1112 | 55.6 | 4966 | 24 | AAD37256 | Adeno-associated v |
| 20 | 1112 | 55.6 | 4990 | 24 | AAD37262 | Adeno-associated v |
| 21 | 1002.6 | 50.1 | 1821 | 24 | AAD37241 | Human dystrophin r |
| 22 | 1002.6 | 50.1 | 2169 | 24 | AAD37232 | Human dystrophin r |
| 23 | 1002.6 | 50.1 | 5952 | 22 | AAD06794 | Human dystrophin g |
| 24 | 1002.6 | 50.1 | 8689 | 24 | ABK82000 | DNA encoding mini- |
| 25 | 1002.6 | 50.1 | 11058 | 24 | AAD37229 | Human dystrophin p |
| 26 | 1002.6 | 50.1 | 11241 | 24 | ABK82005 | cDNA encoding huma |
| 27 | 1002.6 | 50.1 | 11443 | 24 | ABK82002 | DNA encoding mini- |
| 28 | 1002.6 | 50.1 | 12923 | 10 | AAN90338 | Sequence of human |
| 29 | 1002.6 | 50.1 | 13957 | 24 | ABT10904 | Human breast cance |
| 30 | 1002.6 | 50.1 | 13957 | 24 | ABG69900 | Human dystrophin g |
| 31 | 1002.6 | 50.1 | 13957 | 24 | ABN95786 | Gene #2284 used to |
| 32 | 1002.6 | 50.1 | 13957 | 24 | ABK81959 | cDNA encoding huma |
| 33 | 1001 | 50.0 | 1434 | 24 | AAD37243 | Human dystrophin r |
| 34 | 991.6 | 49.6 | 13977 | 24 | ABS70403 | Human bone remodel |
| 35 | 985.8 | 49.3 | 1991 | 24 | AAD37231 | Human dystrophin N |
| 36 | 869.8 | 43.5 | 13815 | 24 | ABK81960 | cDNA encoding mous |
| 37 | 869.8 | 43.5 | 13815 | 24 | ABI99799 | Mouse ischaemic co |
| 38 | 869.8 | 43.5 | 19307 | 17 | AAT27558 | Shuttle vector pAd |
| 39 | 868.2 | 43.4 | 13815 | 19 | AAV18885 | Mus musculus dyst |
| 40 | 767 | 38.3 | 3275 | 10 | AAN97129 | Partial sequence o |
| 41 | 735.4 | 36.8 | 4075 | 21 | AAZ48569 | A rod shortened dy |
| 42 | 731.2 | 36.5 | 3747 | 21 | AAZ48566 | A rod shortened dy |
| 43 | 677.6 | 33.9 | 3163 | 21 | AAZ48571 | A rod shortened dy |
| 44 | 665 | 33.2 | 1667 | 24 | AAD37235 | Human dystrophin N |
| 45 | 567.4 | 28.4 | 2005 | 25 | ABT33376 | NOVX DNA sequence |

ALIGNMENTS

RESULT 1
AAD37242
ID AAD37242 standard; DNA; 3446 BP.
XX
AC AAD37242;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta3447.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular

QY 1801 CTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGAAGACAAGTACAGATACC 1860
Dbb|||||
2800 CTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGAAGACAAGTACAGATACC 2859
QY 1861 TTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGCTGGGCCTCCTTC 1920
Dbb|||||
2860 TTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGCTGGGCCTCCTTC 2919
QY 1921 TGCATGATTCTATCCAAATTCACAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGCAGTA 1980
Dbb|||||
2920 TGCATGATTCTATCCAAATTCACAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGCAGTA 2979
QY 1981 ACATTGAGCCCAAGTGTCCGGA 2001
Dbb|||||
2980 ACATTGAGCCCAAGTGTCCGGA 3000
RESULT 2
AAD37260
ID AAD37260 standard; DNA; 4414 BP.
XX
AC AAD37260;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-3447.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Unidentified.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 65-66; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence.
XX
SQ Sequence 4414 BP; 1255 A; 1075 C; 1086 G; 998 T; 0 other;

Query Match 99.5%; Score 1990; DB 24; Length 4414;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA 60
Dbb|||||
1757 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA 1816
QY 61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAAT 120
Dbb|||||
1817 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAAT 1876
QY 121 GATGTGGAAGTGGTGAAGACCAAGTTTCATCTACTCATGAGGGGTACATGATGGATTTGACA 180
Dbb|||||
1877 GATGTGGAAGTGGTGAAGACCAAGTTTCATCTACTCATGAGGGGTACATGATGGATTTGACA 1936
QY 181 GCCCATCAGGGCCGGTGGTAATATTTCTACAATTTGGGAAGTAAGCTGATTTGGAACAGGA 240
Dbb|||||
1937 GCCCATCAGGGCCGGTGGTAATATTTCTACAATTTGGGAAGTAAGCTGATTTGGAACAGGA 1996
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCTTAAATTCAGA 300
Dbb|||||
1997 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCTTAAATTCAGA 2056
QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTTACATAGATTTTAAATG 360
Dbb|||||
2057 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTTACATAGATTTTAAATG 2116
QY 361 GATCTCCAGAAATC-GAAACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGA 419
Dbb|||||
2117 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGA 2176
QY 420 ACAAGGAAATGGAGGAAGAGCCTCTTGACCTGATCTTGAAGACCTTAAACCGCCAAGTA 479
Dbb|||||
2177 ACAAGGAAATGGAGGAAGAGCCTCTTGACCTGATCTTGAAGACCTTAAACCGCCAAGTA 2236
QY 480 CAACAACATAAGGTGCTTCAAGAAGATCTAGAAACAAGCAAGTCAGGGTCAATCTCTC 539
Dbb|||||
2237 CAACAACATAAGGTGCTTCAAGAAGATCTAGAAACAAGCAAGTCAGGGTCAATCTCTC 2296
QY 540 ACTCACATGGTGGTGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 599
Dbb|||||
2297 ACTCACATGGTGGTGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 2356
QY 600 GAACAACCTTAAGTATTTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGC 659
Dbb|||||
2357 GAACAACCTTAAGTATTTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGC 2416
QY 660 TGGGTTCTTTTACAAGACATCCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTT 719
Dbb|||||
2417 TGGGTTCTTTTACAAGACATCCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTT 2476
QY 720 TTTAGTGCATGGCTTTTCAGAAAAAGAGATGCAGTGAACAAGATTCACACAACACTGGCTTT 779
Dbb|||||
2477 TTTAGTGCATGGCTTTTCAGAAAAAGAGATGCAGTGAACAAGATTCACACAACACTGGCTTT 2536
QY 780 AAAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACCTGGCCGTTTTAAAAGCGGATCTA 839
Dbb|||||
2537 AAAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACCTGGCCGTTTTAAAAGCGGATCTA 2596
QY 840 GAAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAACAAGATCTTTCTTCAACA 899
Dbb|||||
2597 GAAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAACAAGATCTTTCTTCAACA 2656
QY 900 CTGAAGAATAAGTCAGTGACCCAGAGACGGGAAGCATGGCTGGATAAATTTGCCCGGTGT 959
Dbb|||||
2657 CTGAAGAATAAGTCAGTGACCCAGAGACGGGAAGCATGGCTGGATAAATTTGCCCGGTGT 2716
QY 960 TGGGATAATTTAGTCCAAAAACTTGAAAAAGAGTACAGCACAGACCCCTTGAAAAGACTCCAG 1019
Dbb|||||
2717 TGGGATAATTTAGTCCAAAAACTTGAAAAAGAGTACAGCACAGACCCCTTGAAAAGACTCCAG 2776
QY 1020 GAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAG 1079
Dbb|||||
2777 GAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAG 2836

QY 1080 GGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCAAGATCACCTCGAGAAA 1139
Db |||||
QY 2837 GGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTCAAGATCACCTCGAGAAA 2896
Db |||||
QY 1140 GTCAAGGCACCTTCGAGGAGAAATTGCGCTCTGAAAGAGAACGTGAGCCACGTCAATGAC 1199
Db |||||
QY 2897 GTCAAGGCACCTTCGAGGAGAAATTGCGCTCTGAAAGAGAACGTGAGCCACGTCAATGAC 2956
Db |||||
QY 1200 CTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCACTCTG 1259
Db |||||
QY 2957 CTTGCTCGCCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCACTCTG 3016
Db |||||
QY 1260 GAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGSCAG 1319
Db |||||
QY 3017 GAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGSCAG 3076
Db |||||
QY 1320 CTGCATGAAGCCACACAGGACTTTGGTCCAGCATCTCAGCAGCTTTCTTTCCACGTCTGTC 1379
Db |||||
QY 3077 CTGCATGAAGCCACACAGGACTTTGGTCCAGCATCTCAGCAGCTTTCTTTCCACGTCTGTC 3136
Db |||||
QY 1380 CAGGCTCCCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAG 1439
Db |||||
QY 3137 CAGGCTCCCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAG 3196
Db |||||
QY 1440 ACTCAAACAACCTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGAC 1499
Db |||||
QY 3197 ACTCAAACAACCTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGAC 3256
Db |||||
QY 1500 CTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCAGAAG 1559
Db |||||
QY 3257 CTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCAGAAG 3316
Db |||||
QY 1560 GCCCTTTGCTGGATCTCTTGAGCCCTGTGAGCTGCATGTATGCCCTTGACCAGCACAAC 1619
Db |||||
QY 3317 GCCCTTTGCTGGATCTCTTGAGCCCTGTGAGCTGCATGTATGCCCTTGACCAGCACAAC 3376
Db |||||
QY 1620 CTCAAGCAAAATGACACGCCCCATGGATATCCTGCAGATTATTAATGTTGACCACTATT 1679
Db |||||
QY 3377 CTCAAGCAAAATGACACGCCCCATGGATATCCTGCAGATTATTAATGTTGACCACTATT 3436
Db |||||
QY 1680 TATGACCGCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTCGTGGATATG 1739
Db |||||
QY 3437 TATGACCGCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTCGTGGATATG 3496
Db |||||
QY 1740 TGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGTCTG 1799
Db |||||
QY 3497 TGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGTCTG 3556
Db |||||
QY 1800 TCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGATAC 1859
Db |||||
QY 3557 TCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGATAC 3616
Db |||||
QY 1860 CTTTTCAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCCAGGCGAGGCTGGGCCCTCTT 1919
Db |||||
QY 3617 CTTTTCAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCCAGGCGAGGCTGGGCCCTCTT 3676
Db |||||
QY 1920 CTGCATGATTCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGCAGT 1979
Db |||||
QY 3677 CTGCATGATTCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGCAGT 3736
Db |||||
QY 1980 AACATTGAGCCCAAGTGTCCGGA 2001
Db |||||
QY 3737 AACATTGAGCCCAAGTGTCCGGA 3758
Db |||||

RESULT 3
ABK81997
ID ABK81997 standard; DNA; 5417 BP.
XX
AC ABK81997;
XX
DT 13-AUG-2002 (first entry)
XX

DE DNA encoding mini-dystrophin protein deltaR4-R23.
XX
KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.
OS Homo sapiens.
OS Synthetic.
XX
PN WO200229056-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US31126.
XX
PR 06-OCT-2000; 2000US-238848P.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Chamberlain JS, Harper SQ;
XX
DR WPI; 2002-435334/46.
XX
PT A composition for preparing therapeutic drugs, has a mini-dystrophin
PT peptide comprising a specific number of spectrin-like repeat domains,
PT or a nucleic acid sequence encoding the mini-dystrophin peptide -
XX
PS Disclosure; Fig 12; 145pp; English.
XX
CC The invention describes a composition comprising a mini-dystrophin
CC peptide comprising a spectrin-like repeat domain, where the domain
CC comprises n spectrin-like repeats, and contains no more than n
CC spectrin-like repeats, where n is an even number between 4-24, or a
CC nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin
CC peptide or the polynucleotide encoding it is useful as a medicament,
CC for preparing a drug for therapeutic application and in the preparation
CC of a composition for treatment of muscle disease, e.g. Duchenne's
CC muscular dystrophy (DMD). This sequence represents a mini-dystrophin
CC sequence of the invention.
XX
SQ Sequence 5417 BP; 1700 A; 1192 C; 1182 G; 1343 T; 0 other;

Query Match 91.0%; Score 1820.8; DB 24; Length 5417;
Best Local Similarity 92.6%; Pred.No. 0;
Matches 1999; Conservative 0; Mismatches 2; Indels 157; Gaps 2;

QY 1 GGCAGTTCAATTGATGGAGAGTGAGTAACCTGGACCGTTATCAACAGCTTTAGAAGAA 60
Db |||||
QY 1199 GGCAGTTCAATTGATGGAGAGTGAGTAACCTGGACCGTTATCAACAGCTTTAGAAGAA 1258
Db |||||
QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAAT 120
Db |||||
QY 1259 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAAT 1318
Db |||||
QY 121 GATGTGGAAGTGGTGAAGACCCAGTTTCATCTCATGAGGGGTACATGATGGATTGACA 180
Db |||||
QY 1319 GATGTGGAAGTGGTGAAGACCCAGTTTCATCTCATGAGGGGTACATGATGGATTGACA 1378
Db |||||
QY 181 GCCCATCAGGCGCGGTTGGTATATTTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240
Db |||||
QY 1379 GCCCATCAGGCGCGGTTGGTATATTTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 1438
Db |||||
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCTAAATTTCAAGA 300
Db |||||
QY 1439 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCTAAATTTCAAGA 1498
Db |||||
QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGATTTTAATG 360
Db |||||
QY 1499 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGATTTTAATG 1558
Db |||||
QY 361 GATCTCCAGATC-GAAACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAGAAAGA 419
Db |||||
QY 1559 GATCTCCAGATCAGAAACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAGAAAGA 1618
Db |||||

QY 420 ACACGAGAAATGGAGGAAGAGCCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTA 479
Db 1619 ACACGAGAAATGGAGGAAGAGCCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTA 1678
QY 480 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACACGAACAAGTCAGGGTCAATTCTCTC 539
Db 1679 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACACGAACAAGTCAGGGTCAATTCTCTC 1738
QY 540 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 599
Db 1739 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 1798
QY 600 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACAACATCTGTAGATGGACAGACCCGC 659
Db 1799 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACAACATCTGTAGATGGACAGACCCGC 1858
QY 660 TGGGTTCTTTTACAAGACATCCTTCTCAAAATGGCAACGTCTTACTGAAGAACAGTGCCTT 719
Db 1859 TGGGTTCTTTTACAAGACATCCTTCTCAAAATGGCAACGTCTTACTGAAGAACAGTGCCTT 1918
QY 720 TTTAGTGCATGGCTTTCAAGAAAAGAGATGCAAGTGAACAAGATTCACAACTGGCTTT 779
Db 1919 TTTAGTGCATGGCTTTCAAGAAAAGAGATGCAAGTGAACAAGATTCACAACTGGCTTT 1978
QY 780 AAAGATCAAAATGAATGTTATCAAGTCTTCAAAAACCTGGCCGTTTTAAAGCGGATCTA 839
Db 1979 AAAGATCAAAATGAATGTTATCAAGTCTTCAAAAACCTGGCCGTTTTAAAGCGGATCTA 2038
QY 840 GAAAAGAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACA 899
Db 2039 GAAAAGAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACA 2098
QY 900 CTGAAGAATAAGTCAGTACCCAGAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGT 959
Db 2099 CTGAAGAATAAGTCAGTACCCAGAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGT 2158
QY 960 TGGGATAATTTAGTCAAAAAACTTGAAAAGAGTACAGCACAGA----- 1002
Db 2159 TGGGATAATTTAGTCAAAAAACTTGAAAAGAGTACAGCACAGATTTCACAGGCTGTACC 2218
QY 1003 ----- 1002
Db 2219 ACCACTCAGCCATCACTAACACAGACAACACTGTAATGGAAACAGTAACACTACGGTGACCACA 2278
QY 1003 ----- 1002
Db 2279 AGGGAACAGATCCTGGTAAAGCATGCTCAAGAGGAACTTCCACCACCACCTCCCCAAAAAG 2338
QY 1003 -----CCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACCGGATGAG 1043
Db 2339 AAGAGGCAGATTACTGTGGATCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACCGGATGAG 2398
QY 1044 CTGGACCTCAAGTCGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGAT 1103
Db 2399 CTGGACCTCAAGTCGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGAT 2458
QY 1104 CTCCTCATTTGACTCTCTCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATT 1163
Db 2459 CTCCTCATTTGACTCTCTCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATT 2518
QY 1164 GCGCCTCTGAAAGAGAACGCTGAGCCACGCTCAATGACCTTGTCTGCCAGCTTACCACCTTG 1223
Db 2519 GCGCCTCTGAAAGAGAACGCTGAGCCACGCTCAATGACCTTGTCTGCCAGCTTACCACCTTG 2578
QY 1224 GGCATTTCAGCTCTCACCGGTATACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAG 1283
Db 2579 GGCATTTCAGCTCTCACCGGTATACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAG 2638
QY 1284 CTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCCACAGGACTTT 1343
Db 2639 CTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCCACAGGACTTT 2698
QY 1344 GGTCCAGCATCTCAGCACTTTCTTTCCACGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATC 1403

Db 2699 GGTCCAGCATCTCAGCACTTTCTTTCCACGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATC 2758
QY 1404 TCGCAAAACAAAGTCCCTACTATATCAACCACGAGACTCAAAACAATTGCTGGACCAT 1463
Db 2759 TCGCAAAACAAAGTCCCTACTATATCAACCACGAGACTCAAAACAATTGCTGGACCAT 2818
QY 1464 CCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCT 1523
Db 2819 CCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCT 2878
QY 1524 TATAGGACTGCCATGAAACTCCGAAGACTGCAGAAAGCCCTTTGCTTGGATCTCTTGAGC 1583
Db 2879 TATAGGACTGCCATGAAACTCCGAAGACTGCAGAAAGCCCTTTGCTTGGATCTCTTGAGC 2938
QY 1584 CTGTCAGCTGCATGTGATGCCCTTGGACCAAGCAACCTCAAGCAAAATGACCAAGCCCATG 1643
Db 2939 CTGTCAGCTGCATGTGATGCCCTTGGACCAAGCAACCTCAAGCAAAATGACCAAGCCCATG 2998
QY 1644 GATATCCTGCAGATTATTAATTGTTTGAACCACTATTTATGACCGCCTGGAGCAAGAGCAC 1703
Db 2999 GATATCCTGCAGATTATTAATTGTTTGAACCACTATTTATGACCGCCTGGAGCAAGAGCAC 3058
QY 1704 AACAAATTTGGTCAACCTCCCTCTCTCGCTGGATATGTCTGAACTGGCTGCTGAATGTT 1763
Db 3059 AACAAATTTGGTCAACCTCCCTCTCTCGCTGGATATGTCTGAACTGGCTGCTGAATGTT 3118
QY 1764 TATGATACGGGACGAACAGGGAGGATCCGCTGCTCTTTTAAACTGGCATATTTCC 1823
Db 3119 TATGATACGGGACGAACAGGGAGGATCCGCTGCTCTTTTAAACTGGCATATTTCC 3178
QY 1824 CTGTGTAAAGCACATTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCA 1883
Db 3179 CTGTGTAAAGCACATTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCA 3238
QY 1884 ACAGGATTTTGTGACCAGCGCAGGCTGGGCTCCTTCTGATGATTTATCCAAATTTCCA 1943
Db 3239 ACAGGATTTTGTGACCAGCGCAGGCTGGGCTCCTTCTGATGATTTATCCAAATTTCCA 3298
QY 1944 AGACAGTTGGGTGAAGTTGCACTCCTTTTGGGGCAGTAACATTGAGCCAAAGTGTCCGGA 2001
Db 3299 AGACAGTTGGGTGAAGTTGCACTCCTTTTGGGGCAGTAACATTGAGCCAAAGTGTCCGGA 3356

RESULT 4

AAD37240
ID AAD37240 standard; DNA; 3510 BP.
XX AC AAD37240;
XX DT 21-AUG-2002 (first entry)
XX Human dystrophin minigene delta3510.
DE Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX OS Homo sapiens.
XX WO200183695-A2.
PN 08-NOV-2001.
XX 27-APR-2001; 2001WO-US13677.
XX 28-APR-2000; 2000US-200777P.
XX (XIAO/) XIAO X.
PA Xiao X;
XX WPI; 2002-049342/06.

XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 51-52; 71pp; English.
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3510 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1 and R2), 8407-10227 (rods R23 and R24, hinge H4 and
CC CR domain) and 11047-11058 (dystrophin last 3 amino acids).
XX
SQ Sequence 3510 BP; 1073 A; 787 C; 828 G; 822 T; 0 other;

Query Match 79.6%; Score 1593; DB 24; Length 3510;
Best Local Similarity 87.7%; Pred. No. 0;
Matches 1811; Conservative 0; Mismatches 190; Indels 64; Gaps 4;

QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAAGAA 60
Db |||||
QY 1000 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAAGAA 1059
Db |||||

QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT 120
Db |||||
QY 1060 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT 1119
Db |||||

QY 121 GATGTGGAAGTGTGAAAGACCAGTTTCTACTCATGAGGGGTACATGATGGATTGACA 180
Db |||||
QY 1120 GATGTGGAAGTGTGAAAGACCAGTTTCTACTCATGAGGGGTACATGATGGATTGACA 1179
Db |||||

QY 181 GCCCATCAGGCGCGGTTGGTAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGA 240
Db |||||
QY 1180 GCCCATCAGGCGCGGTTGGTAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGA 1239
Db |||||

QY 241 AAATTATCAGAAGATGAAGAAACCTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGA 300
Db |||||
QY 1240 AAATTATCAGAAGATGAAGAAACCTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGA 1299
Db |||||

QY 301 TGGGAATGCCTCAGGCTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 360
Db |||||
QY 1300 TGGGAATGCCTCAGGCTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 1359
Db |||||

QY 361 GATCTCCAGAATC-GAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGA 419
Db |||||
QY 1360 GATCTCCAGAATCAGAAGTGAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGA 1419
Db |||||

QY 420 ACAAGGAAAAATGGAGGAAGAGCCTCTTTGGACCTGATCTTTGAAGACCTAAACCGCCAAGTA 479
Db |||||
QY 1420 ACAAGGAAAAATGGAGGAAGAGCCTCTTTGGACCTGATCTTTGAAGACCTAAACCGCCAAGTA 1479
Db |||||

QY 480 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACCAAGAACAGTCAAGGTCAATTTCTCTC 539
Db |||||
QY 1480 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACCAAGAACAGTCAAGGTCAATTTCTCTC 1539
Db |||||

QY 540 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAA 599
Db |||||
QY 1540 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAA 1599
Db |||||

QY 600 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGC 659
Db |||||
QY 1600 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGC 1659
Db |||||

QY 660 TGGGTTCTTTTACAAGACATCCTTCTCAATGGCAACGTCCTTACTGAAGAACAGTGCCTT 719
Db |||||

Db |||||
QY 1660 TGGGTTCTTTTACAAGACAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGAA 1719
Db |||||
QY 720 TTTAGTGCATGGCTTTCAGAAAAAGAGATGCAGTGAACAAGATTACACAACCTGGCTTT 779
Db |||||
QY 1720 CTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGC 1779
Db |||||
QY 780 AAAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACCTGGCCGTTTTAAAAAGCGGATCTA 839
Db |||||
QY 1780 GACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAAA 1839
Db |||||
QY 840 GAAAAAGAAAAAGCAATCCATGGGCAAACTGTA-----TTCACTCAAAACAAG 885
Db |||||
QY 1840 ACTAAAGAACCTGTAATCATGATGACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAG 1899
Db |||||
QY 886 ATCTTTCTTTCAACACTGAAGAATAAGTCAAGTACCCAGAGACCCAGAGACATGGCTGGA-- 943
Db |||||
QY 1900 CCTTTGGAAGGACTAGAGAAAACCTCTACCAAGAGCCAGAGAGCTGCCTCCTGAGGAGAGA 1959
Db |||||
QY 944 -----TAACTTTGCCCGTGTGGGATAAATTTAGTCCAAA 978
Db |||||
QY 1960 GCCCAGAAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAA 2019
Db |||||
QY 979 AACTTTGAA-----AAGAGTACAGCACAGACCCCTTGAAGAGACTC 1016
Db |||||
QY 2020 AAATTGAACCTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGAGACTC 2079
Db |||||
QY 1017 CAGGAACTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATC 1076
Db |||||
QY 2080 CAGGAACTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATC 2139
Db |||||
QY 1077 AAGGGATCCTGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCTCAAGATCACCTCGAG 1136
Db |||||
QY 2140 AAGGGATCCTGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCTCAAGATCACCTCGAG 2199
Db |||||
QY 1137 AAAGTCAAGGCACCTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGCTGAGCCACGTCAT 1196
Db |||||
QY 2200 AAAGTCAAGGCACCTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGCTGAGCCACGTCAT 2259
Db |||||
QY 1197 GACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTGAGCTCTCCTCCTGATAACCTCAGCACT 1256
Db |||||
QY 2260 GACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTGAGCTCTCCTCCTGATAACCTCAGCACT 2319
Db |||||
QY 1257 CTGGAAGACCTGAACACCCAGATGGAAGCTTTGCGAGCTTGGCCGCTCGAGGACCGAGTCAGG 1316
Db |||||
QY 2320 CTGGAAGACCTGAACACCCAGATGGAAGCTTTGCGAGCTTGGCCGCTCGAGGACCGAGTCAGG 2379
Db |||||
QY 1317 CAGCTGCATGAAGCCCAACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCT 1376
Db |||||
QY 2380 CAGCTGCATGAAGCCCAACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCT 2439
Db |||||
QY 1377 GTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCCAACAAAGTGCCTACTATATCAACCAC 1436
Db |||||
QY 2440 GTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCCAACAAAGTGCCTACTATATCAACCAC 2499
Db |||||
QY 1437 GAGACTCAAAACAACCTTGTCTGGGACCATCCCAAAATGACAGAGCTTACCAGTCTTTAGCT 1496
Db |||||
QY 2500 GAGACTCAAAACAACCTTGTCTGGGACCATCCCAAAATGACAGAGCTTACCAGTCTTTAGCT 2559
Db |||||
QY 1497 GACCTGAATAATGT CAGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAG 1556
Db |||||
QY 2560 GACCTGAATAATGT CAGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAG 2619
Db |||||
QY 1557 AAGGCCCTTTGCTTGGATCTCTTGAGCCTGT CAGCTGCTATGATGCTTGGACCGACAC 1616
Db |||||
QY 2620 AAGGCCCTTTGCTTGGATCTCTTGAGCCTGT CAGCTGCTATGATGCTTGGACCGACAC 2679
Db |||||
QY 1617 AACCTCAAGCAAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACT 1676
Db |||||
QY 2680 AACCTCAAGCAAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACT 2739
Db |||||
QY 1677 ATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGGAT 1736
Db |||||

| | | | |
|----|------|--|------|
| Db | 2740 | ATTTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCCCTCTCTGGTGGAT | 2799 |
| QY | 1737 | ATGTGTCGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTC | 1796 |
| Db | 2800 | ATGTGTCGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTC | 2859 |
| QY | 1797 | CTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGA | 1856 |
| Db | 2860 | CTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGA | 2919 |
| QY | 1857 | TACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGCAGGTGGGCCTC | 1916 |
| Db | 2920 | TACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGCAGGTGGGCCTC | 2979 |
| QY | 1917 | CTTCTGCATGATCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGC | 1976 |
| Db | 2980 | CTTCTGCATGATCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGC | 3039 |
| QY | 1977 | AGTAACATTTGACCAAGTGTCCGGA | 2001 |
| Db | 3040 | AGTAACATTTGACCAAGTGTCCGGA | 3064 |

RESULT 5
AAD37259
ID AAD37259 standard; DNA; 4476 BP.
XX
AC AAD37259;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-3510.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.

New dystrophin minigene for treating Duchenne or Becker muscular dystrophy comprises an N-terminal domain or modified N-terminal domain, rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene -

PS Example 1; Page 63-65; 71pp; English.

The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is AAV vector plasmid construct containing human dystrophin minigenes, a muscle creatine kinase (MCK) promoter and a small polyA signal sequence.

Db 2716 GCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACCTGACTGGAA 2775

QY 979 AACTTGAA-----AAGAGTACAGCACAGACCCCTTGAAAGACTC 1016

Db 2776 AAATTGAACCTGCACCTCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAAGACTC 2835

QY 1017 CAGGAACCTCAAGAGGCCACCGATGAGCTGGACCTCAAGTCGCGCAAGCTGAGGTGATC 1076

Db 2836 CAGGAACCTCAAGAGGCCACCGATGAGCTGGACCTCAAGTCGCGCAAGCTGAGGTGATC 2895

QY 1077 AAGGATCCTGGCAGCCCGTGGCGCATCTCCTCATTTGAGTCTCTCCAAGATCACTCGAG 1136

Db 2896 AAGGATCCTGGCAGCCCGTGGCGCATCTCCTCATTTGAGTCTCTCCAAGATCACTCGAG 2955

QY 1137 AAAGTCAAGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTCGAGCCACGTCAAT 1196

Db 2956 AAAGTCAAGCACCTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTCGAGCCACGTCAAT 3015

QY 1197 GACCTTGCTCGCCAGCTTACCATTGCGCATTCAGCTCTCACCGTATAACCTCAGCACT 1256

Db 3016 GACCTTGCTCGCCAGCTTACCATTGCGCATTCAGCTCTCACCGTATAACCTCAGCACT 3075

QY 1257 CTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCGTCGAGGACCGAGTCAGG 1316

Db 3076 CTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCGTCGAGGACCGAGTCAGG 3135

QY 1317 CAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTCCACGTCT 1376

Db 3136 CAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTCCACGTCT 3195

QY 1377 GTCCAGGGTCCCTGGGAGAGAGCCATCTGCCAAACAAAGTGCCTACTATATCAACCAC 1436

Db 3196 GTCCAGGGTCCCTGGGAGAGAGCCATCTGCCAAACAAAGTGCCTACTATATCAACCAC 3255

QY 1437 GAGACTCAACAACTTGCTGGGACCATCCCAAATGACAGAGCTTACCAGTCTTTAGCT 1496

Db 3256 GAGACTCAACAACTTGCTGGGACCATCCCAAATGACAGAGCTTACCAGTCTTTAGCT 3315

QY 1497 GACCTGAATAATGTGAGTTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCAG 1556

Db 3316 GACCTGAATAATGTGAGTTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCAG 3375

QY 1557 AAGGCCCTTTGCTGGATCTCTTGAGCCTGTGAGTGCATGTGATGCCCTTGGACCCAGC 1616

Db 3376 AAGGCCCTTTGCTGGATCTCTTGAGCCTGTGAGTGCATGTGATGCCCTTGGACCCAGC 3435

QY 1617 AACCTCAAGCAAAATGACAGCCCATGATATCCTGCAGATTATTAATTGTTTGACCACT 1676

Db 3436 AACCTCAAGCAAAATGACAGCCCATGATATCCTGCAGATTATTAATTGTTTGACCACT 3495

QY 1677 ATTTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTCGCTGGAT 1736

Db 3496 ATTTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTCGCTGGAT 3555

QY 1737 ATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGTC 1796

Db 3556 ATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGTC 3615

QY 1797 CTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGA 1856

Db 3616 CTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGA 3675

QY 1857 TACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGGCCTC 1916

Db 3676 TACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGGCCTC 3735

QY 1917 CTTCTGCATGATTTCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGC 1976

Db 3736 CTTCTGCATGATTTCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGC 3795

QY 1977 AGTAACATTGAGCCCAAGTGTCCGGA 2001

Db 3796 AGTAACATTGAGCCCAAGTGTCCGGA 3820

RESULT 6

AAZ48568
ID AAZ48568 standard; cDNA to mRNA; 4402 BP.

XX
AC AAZ48568;

XX
DT 31-MAR-2000 (first entry)

XX
PN JP11318467-A.

XX
PD 24-NOV-1999.

XX
PF 08-MAY-1998; 98JP-0142134.

XX
PR 08-MAY-1998; 98JP-0142134.

XX
PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.

XX
PA (KOKU-) KOKURITSU SEISHIN SHINKAI CENT SOCHO.

XX
DR WPI; 2000-100771/09.

XX
DR P-PSDB; AAY59239.

XX
PT A gene encoding a shortened dystrophin - useful for the treatment of

XX
XX muscular dystrophy

PS Claim 7; Page 21-22; 44pp; Japanese.

CC The invention provides a gene for the treatment of muscular dystrophy

CC having at least one rod repeat structure of hinge 1, hinge 4 and rod

CC domain of dystrophin gene and having a base sequence of 4.5 kb. The gene

CC and a gene-introducing medium consisting of an adeno-associated virus

CC (AAV) vector or lentivirus vector containing the rod shortened dystrophin

CC genes can be used for the genetic treatment of muscular dystrophy of low

CC immune reaction. The present sequence represents a rod shortened

CC dystrophin encoding sequence.

XX Sequence 4402 BP; 1339 A; 984 C; 1010 G; 1069 T; 0 other;

QY Query Match 66.6%; Score 1332; DB 21; Length 4402;

Db Best Local Similarity 83.8%; Pred. No. 0;

QY Matches 1677; Conservative 0; Mismatches 0; Indels 325; Gaps 2;

QY 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAA 60

Db 1195 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAA 1254

QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAGAGATTTCCTAAT 120

Db 1255 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAGAGATTTCCTAAT 1314

QY 121 GATGTGGAAGTGGTGAAGACCCAGTTTCATACTCATGAGGGTACATGATGATTGACA 180

Db 1315 GATGTGGAAGTGGTGAAGACCCAGTTTCATACTCATGAGGGTACATGATGATTGACA 1374

QY 181 GCCCATCAGGCGGGTTGGTAAATATTTACAATTGGGAAGTAAAGCTGATTGGAACAGGA 240

Db 1375 GCCCATCAGGCGGGTTGGTAAATATTTACAATTGGGAAGTAAAGCTGATTGGAACAGGA 1434

QY 241 AAATTATCAGAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGA 300

Db 1435 AAATTATCAGAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGA 1494

QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGGAAAAACAAAGCAATTTACATAGATTTTAATG 360

Db

Db 1495 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 1554

QY 361 GATCTCCAGAAATC-GAAACTGAAAGAGTTGAAATGACTGGCTAAACAAACAGAAAGAA 419

Db 1555 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAAATGACTGGCTAAACAAACAGAAAGAA 1614

QY 420 ACAAGGAAAAATGGAGAAAGAGCCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 479

Db 1615 ACAAGGAAAAATGGAGAAAGAGCCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 1674

QY 480 CAACAAACATAAGGTGTTCAAGAAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTC 539

Db 1675 CAACAAACATAAGGTGTTCAAGAAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTC 1734

QY 540 ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGTTTGGAA 599

Db 1735 ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGTTTGGAA 1794

QY 600 GAAACAACTTAAGGTATTTGGGAGATCGATGGSCAAACATCTGTAGATGGACAGAACCCGC 659

Db 1795 GAAACAACTTAAGGTATTTGGGAGATCGATGGSCAAACATCTGTAGATGGACAGAACCCGC 1854

QY 660 TGGGTTCTTTTACAGACATCCTTCTCAATGGCAACGCTCTTACTGAAGAACAGTGGCTT 719

Db 1855 TGGGTTCTTTTACAGACATCCTTCTCAATGGCAACGCTCTTACTGAAGAACAGTGGCTT 1914

QY 720 TTTAGTGCAATGGCTTTCAGAAAAAGAAAGATGCAAGTGAACAAGATTCACACAACTGGCTT 779

Db 1915 TTTAGTGCAATGGCTTTCAGAAAAAGAAAGATGCAAGTGAACAAGATTCACACAACTGGCTT 1974

QY 780 AAAGATCAAAATGAATGTTATCAAGTCTTCAAAAACTGGCCGTTTAAAGCGGATCTA 839

Db 1975 AAAGATCAAAATGAATGTTATCAAGT 2001

QY 840 GAAAAGAAAAAGCAATCCATGGGCAACTGTATTCACTCAAAACAAGATCTTCTTTCACA 899

Db 2002 ----- 2001

QY 900 CTGAAGAATAAGTCAGTGACCCAGAAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGT 959

Db 2002 ----- 2001

QY 960 TGGGATAATTTAGTCCAAAAAATTGAAAAAGAGTACAGCACAGACCCCTTGAAAGACTCCAG 1019

Db 2002 ----- 2001

QY 1020 GAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAG 1079

Db 2002 ----- 2001

QY 1080 GGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCAAGATCACCTCGAGAAA 1139

Db 2002 -----CTCGAGAAA 2010

QY 1140 GTCAAGGCACTTCGAGGAGAAATTGCGCTCTGAAAGAGAAAGCTGAGCCACGTCAATGAC 1199

Db 2011 GTCAAGGCACTTCGAGGAGAAATTGCGCTCTGAAAGAGAAAGCTGAGCCACGTCAATGAC 2070

QY 1200 CTTGCTCGCCAGCTTACCACCTTTGGSCATTCAGCTCTCACCGTATAACCTCAGCACTCTG 1259

Db 2071 CTTGCTCGCCAGCTTACCACCTTTGGSCATTCAGCTCTCACCGTATAACCTCAGCACTCTG 2130

QY 1260 GAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTACGGCAG 1319

Db 2131 GAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTACGGCAG 2190

QY 1320 CTGCATGAAGCCCAACAGGACTTTTGTGCCAGCATCTCAGCACTTCTTTCCACGTCTGTC 1379

Db 2191 CTGCATGAAGCCCAACAGGACTTTTGTGCCAGCATCTCAGCACTTCTTTCCACGTCTGTC 2250

QY 1380 CAGGGTCCCTGGGAGAGAGCCATCTCGCAAAACAAAGTGCCCTACTATATCAACCCAGAG 1439

Db 2251 CAGGGTCCCTGGGAGAGAGCCATCTCGCAAAACAAAGTGCCCTACTATATCAACCCAGAG 2310

QY 1440 ACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCGAGCTTTAGCTGAC 1499

Db 2311 ACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCGAGCTTTAGCTGAC 2370

QY 1500 CTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAAACCTCCGAGACTGCAGAAG 1559

Db 2371 CTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAAACCTCCGAGACTGCAGAAG 2430

QY 1560 GCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGACCAGCAACAAC 1619

Db 2431 GCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGACCAGCAACAAC 2490

QY 1620 CTCGAAGCAAAATGACCCCATGGATATCCTGCAGATTATTAATTGTTGACCACCTATT 1679

Db 2491 CTCGAAGCAAAATGACCCCATGGATATCCTGCAGATTATTAATTGTTGACCACCTATT 2550

QY 1680 TATGACCGCCTGGAGCAAGAGACAAACAATTTGGTCAACGTCCCTCTCTGCGTGGATATG 1739

Db 2551 TATGACCGCCTGGAGCAAGAGACAAACAATTTGGTCAACGTCCCTCTCTGCGTGGATATG 2610

QY 1740 TGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCTTG 1799

Db 2611 TGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCTTG 2670

QY 1800 TCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAGATAC 1859

Db 2671 TCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAGATAC 2730

QY 1860 CTTTTCGAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAGCGCAGGCTGGCCCTCCTT 1919

Db 2731 CTTTTCGAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAGCGCAGGCTGGCCCTCCTT 2790

QY 1920 CTGCATGATTTCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGCAGT 1979

Db 2791 CTGCATGATTTCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGCAGT 2850

QY 1980 AACATTGAGCCCAAGTGTCCGGA 2001

Db 2851 AACATTGAGCCCAAGTGTCCGGA 2872

RESULT 7

AAZ48567

ID AAZ48567 standard; cDNA to mRNA; 4402 BP.

AC AAZ48567;

DT 31-MAR-2000 (first entry)

DE A rod shortened dystrophin (deltaDysAx11) encoding nucleotide sequence.

KW Muscular dystrophy; rod domain; adeno-associated virus; AAV;
XX dystrophin gene; truncated; ds.

OS Homo sapiens.

PN JP11318467-A.

PD 24-NOV-1999.

PF 08-MAY-1998; 98JP-0142134.

PR 08-MAY-1998; 98JP-0142134.

PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX (KOKU-) KOKURITSU SEISHIN SHINKEI CENT SOCHO.

DR WPI; 2000-100771/09.

DR P-PSDB; AAY59238.

PT A gene encoding a shortened dystrophin - useful for the treatment of
XX muscular dystrophy

XX Claim 7; Page 16-17; 44pp; Japanese.

CC The invention provides a gene for the treatment of muscular dystrophy

CC having at least one rod repeat structure of hinge 1, hinge 4 and rod

CC domain of dystrophin gene and having a base sequence of 4.5 kb. The gene

CC and a gene-introducing medium consisting of an adeno-associated virus

CC (AAV) vector or lentivirus vector containing the rod shortened dystrophin

CC genes can be used for the genetic treatment of muscular dystrophy of low

CC immune reaction. The present sequence represents a rod shortened

CC dystrophin encoding sequence.

SQ Sequence 4402 BP; 1329 A; 1000 C; 1019 G; 1054 T; 0 other;

Query Match 63.4%; Score 1269.2; DB 21; Length 4402;

Best Local Similarity 82.1%; Pred. No. 0;

Matches 1644; Conservative 0; Mismatches 33; Indels 325; Gaps 3;

QY 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAA 60

Db 1195 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAA 1254

QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTTGCAAGCACACAGGAGAGATTTCTTAAT 120

Db 1255 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTTGCAAGCACACAGGAGAGATTTCTTAAT 1314

QY 121 GATGTGGAAGTGTGAAAGACCAAGTTTCTACTACTCATGAGGGGTACATGATGGATTTGACA 180

Db 1315 GATGTGGAAGTGTGAAAGACCAAGTTTCTACTACTCATGAGGGGTACATGATGGATTTGACA 1374

QY 181 GCCCATCAGGCGCGGTTGGTAATATTTCAAAATTTGGGAAGTAAAGCTGATTGGAACAGGA 240

Db 1375 GCCCATCAGGCGCGGTTGGTAATATTTCAAAATTTGGGAAGTAAAGCTGATTGGAACAGGA 1434

QY 241 AAATTATCAGAAGATGAAGAAACTGAAGTACAAAGAGCAGATGAATCTCTAAATTTCAAGA 300

Db 1435 AAATTATCAGAAGATGAAGAAACTGAAGTACAAAGAGCAGATGAATCTCTAAATTTCAAGA 1494

QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 360

Db 1495 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 1554

QY 361 GATCTCCAGAATC-GAAACTGAAAGAGTTGAATGACTGGCTAAGACCTTAAACGCCAAGTA 419

Db 1555 GATCTCCAGAATCAGAAACTGAAGAGTTGAATGACTGGCTAAGACCTTAAACGCCAAGTA 1614

QY 420 ACAAGGAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 479

Db 1615 ACAAGGAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 1674

QY 480 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAAGTCTCTC 539

Db 1675 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAAGTCTCTC 1734

QY 540 ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 599

Db 1735 ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 1794

QY 600 GAACAACCTTAAGGATTTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAAGACCGC 659

Db 1795 GAACAACCTTAAGGATCTAATACTGAGTGGGAAAA- 1829

QY 660 TGGGTTCTTTTACAAGACATCCTTCTCAAAATGGCAACGTCTTACTGAAGAACAGTGCCTT 719

Db 1830 - 1829

QY 720 TTAGTGATGGCTTTCAGAAAAAGAAAGATGCAGTGAACAAGATTACACAACCTGGCTTT 779

Db 1830 - 1829

QY 780 AAAGATCAAAATGAATGTATCAAGTCTTCAAAAACTGGCCGTTTTAAAAAGCGGATCTA 839

Db 1830 - 1829

QY 840 GAAAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACA 899

Db 1830 - 1829

QY 900 CTGAAGAATAAGTCAGTGACCCAGAGACGGAAGCATGGCTGGATAAATTTTGGCCGGTGT 959

Db 1830 - 1862

QY 960 TGGGATAATTTAGTCCAAAAAATTTGAAAAAGAGTACAGCACAGACCCCTTGAAAGACTCCAG 1019

Db 1863 - 1890

QY 1020 GAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGGCCAAGCTGAGGTGATCAAG 1079

Db 1891 GAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGGCCAAGCTGAGGTGATCAAG 1950

QY 1080 GGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAA 1139

Db 1951 GGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAA 2010

QY 1140 GTCAAGGCACCTTCGAGGAGAAAATTTGGCCTCTGAAAAGAGAACGTGAGCCACGTCAATGAC 1199

Db 2011 GTCAAGGCACCTTCGAGGAGAAAATTTGGCCTCTGAAAAGAGAACGTGAGCCACGTCAATGAC 2070

QY 1200 CTTGCTCGCCAGCTTACCACCTTTGGGCANTCAGCTCTCACCGTATAACCTCAGCACTCTG 1259

Db 2071 CTTGCTCGCCAGCTTACCACCTTTGGGCANTCAGCTCTCACCGTATAACCTCAGCACTCTG 2130

QY 1260 GAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGCAG 1319

Db 2131 GAAGACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGCAG 2190

QY 1320 CTGCATGAAGCCACACAGGACCTTTGGTCCAGCATCTCAGCACCTTTTCCACGTCTGTC 1379

Db 2191 CTGCATGAAGCCACACAGGACCTTTGGTCCAGCATCTCAGCACCTTTTCCACGTCTGTC 2250

QY 1380 CAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCACGAG 1439

Db 2251 CAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCACGAG 2310

QY 1440 ACTCAAAACAACCTTGTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGAC 1499

Db 2311 ACTCAAAACAACCTTGTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGAC 2370

QY 1500 CTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAARCTCCGAAGACTGCAGAAG 1559

Db 2371 CTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAARCTCCGAAGACTGCAGAAG 2430

QY 1560 GCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCCTTGGACCAGCACAAAC 1619

Db 2431 GCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGTGCATGTGATGCCTTGGACCAGCACAAAC 2490

QY 1620 CTCAAGCAAAATGACCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATT 1679

Db 2491 CTCAAGCAAAATGACCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATT 2550

QY 1680 TATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACCGTCCCTCTCTGCGTGGATATG 1739

Db 2551 TATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACCGTCCCTCTCTGCGTGGATATG 2610

QY 1740 TGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCCTG 1799

Db 2611 TGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCCTG 2670

QY 1800 TCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATAC 1859

Db 2671 TCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATAC 2730

QY 1860 CTTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCCAGGCGAGGCTGGGCCCTCCTT 1919

Db 2731 CTTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCCAGGCGAGGCTGGGCCCTCCTT 2790

| | | | | |
|---|---|---|--|------|
| QY | 1920 | CTGCATGATTCATCAAAATTC | CAAGACAGATTGGTGAAAGTTGCATCCTTTGGGGCAGT | 1979 |
| Db | 2791 | CTGCATGATTCATCAAAATTC | CAAGACAGATTGGTGAAAGTTGCATCCTTTGGGGCAGT | 2850 |
| QY | 1980 | AACATTGAGCCCAAGTGTC | CCGGA 2001 | |
| Db | 2851 | AACATTGAGCCCAAGTGTC | CCGGA 2872 | |
| RESULT 8 | | | | |
| ID | AAD37237 | standard; DNA; 3858 BP. | | |
| XX | AAD37237; | | | |
| AC | | | | |
| XX | 21-AUG-2002 | (first entry) | | |
| DT | | | | |
| XX | Human dystrophin minigene | delta3849. | | |
| DE | | | | |
| XX | Human; dystrophin minigene; | muscular; gene therapy; utrophin; spectrin; | | |
| KW | adeno-associated virus; AAV; | Duchenne muscular dystrophy; DMD; BMD; | | |
| KW | Becker muscular dystrophy; | ds. | | |
| XX | | | | |
| OS | Homo. sapiens. | | | |
| XX | WO200183695-A2. | | | |
| PN | | | | |
| XX | 08-NOV-2001. | | | |
| PD | | | | |
| XX | 27-APR-2001; | 2001WO-USI3677. | | |
| PF | | | | |
| XX | 28-APR-2000; | 2000US-200777P. | | |
| PR | | | | |
| XX | (XIAO/) XIAO X. | | | |
| PA | | | | |
| XX | Xiao X; | | | |
| PI | | | | |
| XX | WPI; 2002-049342/06. | | | |
| DR | | | | |
| XX | New dystrophin minigene for treating Duchenne or Becker muscular | | | |
| PT | dystrophy comprises an N-terminal domain or modified N-terminal domain, | | | |
| PT | rod repeats, H1 and H4 domains and a cysteine rich domain of a | | | |
| PT | dystrophin gene - | | | |
| XX | | | | |
| PS | Example 1; Page 48-49; 71pp; English. | | | |
| XX | | | | |
| CC | The present invention relates to an isolated nucleotide sequence encoding | | | |
| CC | a dystrophin minigene. The minigene comprises N-terminal or modified | | | |
| CC | N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 | | | |
| CC | domains and cysteine-rich domains of dystrophin or utrophin genes. The | | | |
| CC | invention also relates to a recombinant adeno-associated virus (AAV) | | | |
| CC | comprising dystrophin minigene operably linked to an expression control | | | |
| CC | element. The dystrophin minigene in operable linkage with an expression | | | |
| CC | useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular | | | |
| CC | dystrophy (BMD) in a mammalian subject. The present sequence is human | | | |
| CC | dystrophin minigene delta3849 containing nucleotides 1-1668 (N-terminus, | | | |
| CC | hinge H1 and rods R1, R2), 8059-10227 (rods R22, R23 and R24, hinge H4 | | | |
| CC | and CR domain) and 11047-11058 (dystrophin last 3 amino acids). | | | |
| XX | | | | |
| SQ | Sequence 3858 BP; 1189 A; 866 C; 905 G; 898 T; 0 other; | | | |
| Query Match 63.1%; Score 1263; DB 24; Length 3858; | | | | |
| Best Local Similarity 75.3%; Pred. No. 0; | | | | |
| Matches 1816; Conservative 0; Mismatches 185; Indels 412; Gaps 3; | | | | |
| QY | 1 | GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAA | 60 | |
| Db | 1000 | GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAA | 1059 | |
| QY | 61 | GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACCAAGGAGAGATTCTTAAT | 120 | |
| Db | 1060 | GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACCAAGGAGAGATTCTTAAT | 1119 | |

QY 1000 ----- 999
Db 2200 GTAATCATGAGTACTTTGAGACTGTACGAATATTCTTGACAGAGCAGCCTTTTGGGAAGGA 2259
QY 1000 ----- 999
Db 2260 CTAGAGAAACTCTACAGAGGCCAGAGAGCTGCCTCCTGAGAGAGAGGCCCAAGATGTC 2319
QY 1000 ----- 999
Db 2320 ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACTG 2379
QY 1000 -----AGACCCCTTGAAGACTCCAGGAACCTTCAA 1028
Db 2380 CACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTCAA 2439
QY 1029 GAGGCCACGGATGAGCTGGACCTCAAGCTCGCCAAAGCTGAGTGATCAAGGGATCCTGG 1088
Db 2440 GAGGCCACGGATGAGCTGGACCTCAAGCTCGCCAAAGCTGAGTGATCAAGGGATCCTGG 2499
QY 1089 CAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCA 1148
Db 2500 CAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCA 2559
QY 1149 CTTGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGTCTGC 1208
Db 2560 CTTGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGTCTGC 2619
QY 1209 CAGCTTACCACTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTG 1268
Db 2620 CAGCTTACCACTTTGGGCATTTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTG 2679
QY 1269 AACACCAGATGGAAGCTTCTGCAAGTGGCCGTCGAGGACCGAGTCAGGCAGCTGATGAA 1328
Db 2680 AACACCAGATGGAAGCTTCTGCAAGTGGCCGTCGAGGACCGAGTCAGGCAGCTGATGAA 2739
QY 1329 GCCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCCTC 1388
Db 2740 GCCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCCTC 2799
QY 1389 TGGGAGAGAGCCATCTGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACA 1448
Db 2800 TGGGAGAGAGCCATCTGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAACA 2859
QY 1449 ACTTGCTGGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGATAAAT 1508
Db 2860 ACTTGCTGGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGATAAAT 2919
QY 1509 GTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGC 1568
Db 2920 GTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGC 2979
QY 1569 TTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCTTGGACAGCACAACTCAAGCAA 1628
Db 2980 TTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCTTGGACAGCACAACTCAAGCAA 3039
QY 1629 AATGACCAGCCCATGGATATCCTGCGAGATTATTAATTGTTTGACCACTATTATGACCGC 1688
Db 3040 AATGACCAGCCCATGGATATCCTGCGAGATTATTAATTGTTTGACCACTATTATGACCGC 3099
QY 1689 CTGGAGCAAGAGCACAAACAATTTGGTCAACGTCCCTCTCTGCGTGGATATGTGTCTGAAC 1748
Db 3100 CTGGAGCAAGAGCACAAACAATTTGGTCAACGTCCCTCTCTGCGTGGATATGTGTCTGAAC 3159
QY 1749 TGGCTGCTGATGTTTATGATACGGGACCAACAGGGAGGATCCGTGCTCTTTTAA 1808
Db 3160 TGGCTGCTGATGTTTATGATACGGGACCAACAGGGAGGATCCGTGCTCTTTTAA 3219
QY 1809 ACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGACAAGTACAGATACCTTTTCAAG 1868
Db 3220 ACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGACAAGTACAGATACCTTTTCAAG 3279
QY 1869 CAAGTGGCAAGTTCACAGGATTTTGTGACCAGCGCAGGCTGGGCGCTCCTTCTGTCATGAT 1928

Db 3280 CAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGCTGGGCTCCTTCTGTCATGAT 3339
QY 1929 TCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGCAGTAACATTGAG 1988
Db 3340 TCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGCAGTAACATTGAG 3399
QY 1989 CCAAGTGTCCGGA 2001
Db 3400 CCAAGTGTCCGGA 3412
RESULT 9
AAD37257
ID AAD37257 standard; DNA; 4825 BP.
XX
AC AAD37257;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-delta3849.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Unidentified.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 61-62; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence.
XX
SQ Sequence 4825 BP; 1369 A; 1175 C; 1204 G; 1077 T; 0 other;

Query Match 63.1%; Score 1263; DB 24; Length 4825;
Best Local Similarity 75.3%; Pred. No. 0;
Matches 1816; Conservative 0; Mismatches 185; Indels 412; Gaps 3;
QY 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTTAGAAGAA 60
Db 1757 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTTAGAAGAA 1816

QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCCTAAT 120
Db 1817 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCCTAAT 1876

QY 121 GATGTGGAAGTGGTGAAGAACCAGTTTTCATACTACTATGAGGGGTACATGATGGATTGACA 180
Db 1877 GATGTGGAAGTGGTGAAGAACCAGTTTTCATACTACTATGAGGGGTACATGATGGATTGACA 1936

QY 181 GCCCATCAGGCCGGGTGGTAATATTTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGA 240
Db 1937 GCCCATCAGGCCGGGTGGTAATATTTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGA 1996

QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGA 300
Db 1997 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGA 2056

QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 360
Db 2057 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 2116

QY 361 GATCTCCAGAAATC-GAAACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGA 419
Db 2117 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGA 2176

QY 420 ACAAGGAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTA 479
Db 2177 ACAAGGAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTA 2236

QY 480 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTTCTTC 539
Db 2237 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTTCTTC 2296

QY 540 ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCAGCGAACTGCTGCTTTGGAA 599
Db 2297 ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCAGCGAACTGCTGCTTTGGAA 2356

QY 600 GAACAACTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGC 659
Db 2357 GAACAACTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGC 2416

QY 660 TGGGTTCTTTTACAAGACATCCCTTCAAAATGGCAACGTCCTTACTGAAGAACAGTGCCTT 719
Db 2417 TGGGTTCTTTTACAAGACATCTCATAGATTACTGCAACAGTTCCCTCCCTGGAAACAG 2476

QY 720 TTTAGTGCAATGGCTTTTACAGAAAAAGAGATGCAGTGAACAAGANTTACACAACTGGCTTT 779
Db 2477 TTTCTTGCCTGGCTTACAGAAGCTGAAAACAACTGCCAATGTCTACAGGATGCTACCCGT 2536

QY 780 A-----AAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACCTGGCCGTT 824
Db 2537 AAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAACAATGGCAAGAC 2596

QY 825 TTAAAGCGGATCTAGAAAAGAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAA 884
Db 2597 CTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAACACGCCAA 2656

QY 885 GATCTTCTTTCAACACTGAAGAATAAGTCAGTGACCCAGAAAGACGGAAGCATGGCTGGAT 944
Db 2657 AAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCCTGTTACAAAGACGTTGGAT 2716

QY 945 AACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAAAACCCTGAAAAGAGTACAGCAC----- 999
Db 2717 AACATGAACCTCAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCCAATTG 2776

QY 1000 ----- 999
Db 2777 GAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGG 2836

QY 1000 ----- 999
Db 2837 CTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGCAGCTTTCCAGCA 2896

QY 1000 ----- 999

Db 2897 GTTCAGAAGCAGAAACGATGTACATAGGGCCTTCAAGAGGGAATTTGAAAAACTAAAGAACCT 2956
QY 1000 ----- 999

Db 2957 GTAATCATGAGTACTCTTTGAGACTGTACCAATATTTCTGACAGAGCAGCCTTTGGAAGGA 3016
QY 1000 ----- 999

Db 3017 CTAGAGAAACTCTACCAGGAGCCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAAATGTC 3076
QY 1000 ----- 999

Db 3077 ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTG 3136
QY 1000 -----AGACCCCTTGAAGACTCCAGGAACCTTCAA 1028

Db 3137 CACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAA 3196
QY 1029 GAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGG 1088
Db 3197 GAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGG 3256
QY 1089 CAGCCCGTGGCGATCTCCTCAATTGACTCTCTCCAAGATCACTCTCGAGAAAGTCAAGGCA 1148
Db 3257 CAGCCCGTGGCGATCTCCTCAATTGACTCTCTCCAAGATCACTCTCGAGAAAGTCAAGGCA 3316
QY 1149 CTTGAGGAGAAATTTGCGCCTCTGAAAAGAGAACGTGAGCCACGTCAATGACCTTGCTGC 1208
Db 3317 CTTGAGGAGAAATTTGCGCCTCTGAAAAGAGAACGTGAGCCACGTCAATGACCTTGCTGC 3376
QY 1209 CAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTTGGAAAGACCTG 1268
Db 3377 CAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTTGGAAAGACCTG 3436
QY 1269 AACACCAGATGGAAGCTTCTGCAAGTGGCCGTGAGGACCGAGTCAAGGACGTGCAATGAA 1328
Db 3437 AACACCAGATGGAAGCTTCTGCAAGTGGCCGTGAGGACCGAGTCAAGGACGTGCAATGAA 3496
QY 1329 GCCCACAGGACTTTGGTCCAGCATCTCAGCACCTTCTTCCACGTCTGTCCAGGGTCCC 1388
Db 3497 GCCCACAGGACTTTGGTCCAGCATCTCAGCACCTTCTTCCACGTCTGTCCAGGGTCCC 3556
QY 1389 TGGGAGAGAGCCCATCTCGCCCAACAAAGTGCCCTACTATATCAACCACGAGACTCAACA 1448
Db 3557 TGGGAGAGAGCCCATCTCGCCCAACAAAGTGCCCTACTATATCAACCACGAGACTCAACA 3616
QY 1449 ACTTGCTGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAAT 1508
Db 3617 ACTTGCTGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAAT 3676
QY 1509 GTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTGC 1568
Db 3677 GTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTGC 3736
QY 1569 TTGGATCTCTTGAGCCTGTGAGTGCATGTGATGCGCTTGGACCAGCAACACCTCAAGCAA 1628
Db 3737 TTGGATCTCTTGAGCCTGTGAGTGCATGTGATGCGCTTGGACCAGCAACACCTCAAGCAA 3796
QY 1629 AATGACCAGCCCATGGATATCCTGCAGATTATTAAATGTTTGACCCACTATTTATGACCGC 1688
Db 3797 AATGACCAGCCCATGGATATCCTGCAGATTATTAAATGTTTGACCCACTATTTATGACCGC 3856
QY 1689 CTGGAGCAAGAGCAACAACAATTTGGTCAACGTCCCTCTCTGCGTGGATATGTGTCTGAAC 1748
Db 3857 CTGGAGCAAGAGCAACAACAATTTGGTCAACGTCCCTCTCTGCGTGGATATGTGTCTGAAC 3916
QY 1749 TGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCTCTTTTAA 1808
Db 3917 TGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCTCTTTTAA 3976
QY 1809 ACTGGCATCATTTCCCTGTGTAAGACACATTTGGAAGACAAGTACAGATACCTTTTCAAG 1868

Db 3977 ACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAGACAAGTACAGATACCTTTTCAAG 4036
QY 1869 CAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGGCTGGGCTCCTTCTGTCATGAT 1928
Db 4037 CAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGGCTGGGCTCCTTCTGTCATGAT 4096
QY 1929 TCTATCCAAATTCOAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGCGAGTAACATTGAG 1988
Db 4097 TCTATCCAAATTCOAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGCGAGTAACATTGAG 4156
QY 1989 CCAAGTGTCCGGA 2001
Db 4157 CCAAGTGTCCGGA 4169

RESULT 10
AAD37263
ID AAD37263 standard; DNA; 4848 BP.
XX
AC AAD37263;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3849.
XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX Chimeric - Homo sapiens.
OS Chimeric - Cytomegalovirus.
OS Chimeric - Unidentified.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -

XX Example 1; Page 68-70; 71pp; English.
PS
XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC cytomegalovirus (CMV) promoter and a small polyA signal sequence.

XX Sequence 4848 BP; 1405 A; 1144 C; 1180 G; 1119 T; 0 other;
PS
XX Query Match 63.1%; Score 1263; DB 24; Length 4848;
XX Best Local Similarity 75.3%; Pred. No. 0;
XX Matches 1816; Conservative 0; Mismatches 185; Indels 412; Gaps 3;

QY 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTTGGACCGTTATCAAAACAGCTTTAGAAGAA 60
Db 1780 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTTGGACCGTTATCAAAACAGCTTTAGAAGAA 1839
QY 61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGCACAAGGAGAGATTCTTAAT 120
Db 1840 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGCACAAGGAGAGATTCTTAAT 1899
QY 121 GATGTGGAAGTGGTGAAGACACAGTTTCTATCTCATGAGGGGTACATGATGATTTGACA 180
Db 1900 GATGTGGAAGTGGTGAAGACACAGTTTCTATCTCATGAGGGGTACATGATGATTTGACA 1959
QY 181 GCCCATCAGGGCCGGGTTGGTAAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240
Db 1960 GCCCATCAGGGCCGGGTTGGTAAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 2019
QY 241 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGACAGATGAATCTCCTAAATTCAAGA 300
Db 2020 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGACAGATGAATCTCCTAAATTCAAGA 2079
QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 360
Db 2080 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 2139
QY 361 GATCTCCAGAATC-GAAACTGAAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAGAAAGA 419
Db 2140 GATCTCCAGAATCAGAATACTGAAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAGAAAGA 2199
QY 420 ACAAGGAAAAATGGAGGAGAGCCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTA 479
Db 2200 ACAAGGAAAAATGGAGGAGAGCCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTA 2259
QY 480 CAACACATTAAGGTGCTTCAAGAAAGATCTAGAACAAAGCAAGTCAGGGTCAATTTCTCTC 539
Db 2260 CAACACATTAAGGTGCTTCAAGAAAGATCTAGAACAAAGCAAGTCAGGGTCAATTTCTCTC 2319
QY 540 ACTCACATGGTGGTGTGATGAATCTTAGTGGAGATCACGCAACTGCTGCTTTGGAA 599
Db 2320 ACTCACATGGTGGTGTGATGAATCTTAGTGGAGATCACGCAACTGCTGCTTTGGAA 2379
QY 600 GAACAACTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAGACCCGC 659
Db 2380 GAACAACTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAGACCCGC 2439
QY 660 TGGGTCTCTTTTACAGACATCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTT 719
Db 2440 TGGGTCTCTTTTACAGACATCTATAGATTACTGCAACAGTTCCCTGGACCTGGAAAG 2499
QY 720 TTTAGTGCATGGCTTTCAGAAAAAGAGATGCAGTGAACAAGATTTCACACACTGGCTTT 779
Db 2500 TTTCTTGCCTGGCTTACAGAACTGAAACAACTGCCAATGTCTTACAGGATGTACCCGT 2559
QY 780 A-----AAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACCTGGCCGTT 824
Db 2560 AAGGAAAGGCTCCTAGAGAGACTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGAC 2619
QY 825 TTAAGAGCGGATCTAGAAAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACA 884
Db 2620 CTCCAAGGTGAAATGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACACGCCAA 2679
QY 885 GATCTTCTTTCAACACTGAAGATAAGTCAGTGACCCAGAAAGCAGGAGCATGGCTGGAT 944
Db 2680 AAAATCCTGAGATCCCTGGAGGTTCCGATGATGCGTCTCTTACAAAGACGTTTGGAT 2739
QY 945 AACTTTGCCCGGTGTTGGGATATTTAGTCCAAAAAAGTGAAGAGAGTACAGCAC----- 999
Db 2740 AACATGAACTTCAAGTGGAGTGAAGTCTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTG 2799
QY 1000 ----- 999
Db 2800 GAAGCCAGTTCTGACCAGTGGAGAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGTGTGG 2859
QY 1000 ----- 999

Db 2860 CTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCA 2919
QY 1000 ----- 999
Db 2920 GTTCAGAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGGAATTGAAAACCTAAAGAACCT 2979
QY 1000 ----- 999
Db 2980 GTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCCTTTGGGAAGGA 3039
QY 1000 ----- 999
Db 3040 CTAGAGAACTCTACCAGGAGCCCGAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAAATGTC 3099
QY 1000 ----- 999
Db 3100 ACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACTGTG 3159
QY 1000 -----AGACCCCTTGAAAGACTCCAGGAACCTCAA 1028
Db 3160 CACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTCAA 3219
QY 1029 GAGGCCACGGATGAGCTGGACCTCAAGCTCGGCCAAGCTGAGGTGATCAAGGGATCCTGG 1088
Db 3220 GAGGCCACGGATGAGCTGGACCTCAAGCTCGGCCAAGCTGAGGTGATCAAGGGATCCTGG 3279
QY 1089 CAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACTCGAGAAAGTCAAGGCA 1148
Db 3280 CAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACTCGAGAAAGTCAAGGCA 3339
QY 1149 CTTTCGAGGAGAAATGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGTCTCGC 1208
Db 3340 CTTTCGAGGAGAAATGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGTCTCGC 3399
QY 1209 CAGCTTACCCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTTGGAAAGACCTG 1268
Db 3400 CAGCTTACCCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTTGGAAAGACCTG 3459
QY 1269 AACACCAGATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAA 1328
Db 3460 AACACCAGATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAA 3519
QY 1329 GCCCACAGGGACTTTGGTCCAGCATCTCAGCACATTCTTTCCACGTCTGTCCAGGGTCCC 1388
Db 3520 GCCCACAGGGACTTTGGTCCAGCATCTCAGCACATTCTTTCCACGTCTGTCCAGGGTCCC 3579
QY 1389 TGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACACGAGACTCAACA 1448
Db 3580 TGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACACGAGACTCAACA 3639
QY 1449 ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAAT 1508
Db 3640 ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAAT 3699
QY 1509 GTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAGGCCCCTTTCG 1568
Db 3700 GTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAGGCCCCTTTCG 3759
QY 1569 TTGGATCTCTTGAGCCTGTGAGCTGATGATGCTTGGACCAAGCAACCTCAAGCAA 1628
Db 3760 TTGGATCTCTTGAGCCTGTGAGCTGATGATGCTTGGACCAAGCAACCTCAAGCAA 3819
QY 1629 AATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTATGACCGC 1688
Db 3820 AATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTATGACCGC 3879
QY 1689 CTGGAGCAAGAGCACAAATTTGGTCAACCTCCCTCTCTGCGTGGATATGTGTCTGAAC 1748
Db 3880 CTGGAGCAAGAGCACAAATTTGGTCAACCTCCCTCTCTGCGTGGATATGTGTCTGAAC 3939
QY 1749 TGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCCGTCTCTTTTAAA 1808

Db 3940 TGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCTCTTTAAA 3999
QY 1809 ACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAG 1868
Db 4000 ACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAG 4059
QY 1869 CAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCTCCTTCTGCATGAT 1928
Db 4060 CAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCTCCTTCTGCATGAT 4119
QY 1929 TCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGCAGTAACATTGAG 1988
Db 4120 TCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGCAGTAACATTGAG 4179
QY 1989 CCAAGTGTCCGGA 2001
Db 4180 CCAAGTGTCCGGA 4192

RESULT 11
AAD37264
ID AAD37264 standard; DNA; 5060 BP.
XX
AC AAD37264;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus (AAV) vector plasmid, AAV-E-CMV-3849.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Cytomegalovirus.
OS Chimeric - Unidentified.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 70-71; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC muscle creatine kinase (MCK) enhancer, a cytomegalovirus (CMV) promoter
CC and a small polyA signal sequence.
XX
SQ Sequence 5060 BP; 1449 A; 1217 C; 1234 G; 1160 T; 0 other;

| | | | | | | | |
|-----------------------|------|---|------|-----------------|-----------------|-------------|--------------|
| Query Match | | | | 63.1%; | Score 1263; | DB 24; | Length 5060; |
| Best Local Similarity | | | | 75.3%; | Pred. No. 0; | | |
| Matches 1816; | | | | Conservative 0; | Mismatches 185; | Indels 412; | Gaps 3; |
| QY | 1 | GGCAGTTCATTGATGGAGAGTGAAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAA | 60 | | | | |
| Db | 1992 | GGCAGTTCATTGATGGAGAGTGAAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAA | 2051 | | | | |
| QY | 61 | GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTTCTAAT | 120 | | | | |
| Db | 2052 | GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTTCTAAT | 2111 | | | | |
| QY | 121 | GATGTGGAAGTGGTGAAGACCAAGTTCATACTCATGAGGGGTACATGATGGATTGACA | 180 | | | | |
| Db | 2112 | GATGTGGAAGTGGTGAAGACCAAGTTCATACTCATGAGGGGTACATGATGGATTGACA | 2171 | | | | |
| QY | 181 | GCCCATCAGGCGCGGTTGGTAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGA | 240 | | | | |
| Db | 2172 | GCCCATCAGGCGCGGTTGGTAATATTCTACAATTGGGAAGTAAAGCTGATTGGAAACAGGA | 2231 | | | | |
| QY | 241 | AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA | 300 | | | | |
| Db | 2232 | AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA | 2291 | | | | |
| QY | 301 | TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATG | 360 | | | | |
| Db | 2292 | TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATG | 2351 | | | | |
| QY | 361 | GATCTCCAGAAATC-GAAACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAA | 419 | | | | |
| Db | 2352 | GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAA | 2411 | | | | |
| QY | 420 | ACAAGGAAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTA | 479 | | | | |
| Db | 2412 | ACAAGGAAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTA | 2471 | | | | |
| QY | 480 | CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTCTCTC | 539 | | | | |
| Db | 2472 | CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTCTCTC | 2531 | | | | |
| QY | 540 | ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA | 599 | | | | |
| Db | 2532 | ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA | 2591 | | | | |
| QY | 600 | GAACAACTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAAACCGC | 659 | | | | |
| Db | 2592 | GAACAACTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAAACCGC | 2651 | | | | |
| QY | 660 | TGGGTTCTTTTACAAGACATCCTTTCTCAAATGGCAACGTTTACTGAAGAACAGTGCCCTT | 719 | | | | |
| Db | 2652 | TGGGTTCTTTTACAAGACACTCATAGATTACTGCAACAGTTCCCTCCCTGGACCTTGAAAG | 2711 | | | | |
| QY | 720 | TTTAGTGCAATGGCTTTTACAGAAAAAGAGATGCAGTGAACAAAGATTCAACAACCTGGCTTT | 779 | | | | |
| Db | 2712 | TTTCTTGCCCTGGCTTACAGAAGCTGAACAACTGCCAATGTCTTACAGGATGCTACCCGT | 2771 | | | | |
| QY | 780 | A-----AAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACTGGCCGTT | 824 | | | | |
| Db | 2772 | AAGGAAAGGCTCCTAGAAAGACTCCAAGGAGTAAAGAGCTGATGAACAAATGGCAAGAC | 2831 | | | | |
| QY | 825 | TTAAAAGCGGATCTAGAAAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAA | 884 | | | | |
| Db | 2832 | CTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAACACAGCCAA | 2891 | | | | |
| QY | 885 | GATCTTCTTTCAACACTGAAGAATAAGTCAAGTGAACCCAGAACGGAAGCATGGCTGGAT | 944 | | | | |
| Db | 2892 | AAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTGTTTACAAAGACGTTTGGAT | 2951 | | | | |
| QY | 945 | AACCTTGGCCCGGTGTGGGATAATTTAGTCCAAAAACTTGAAAAAGAGTACAGCAC----- | 999 | | | | |
| Db | 2952 | AACATGAACTTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTG | 3011 | | | | |
| QY | 1000 | ----- | 999 | | | | |
| Db | 3012 | GAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGTCAGGAACCTTCTGGTGTGG | 3071 | | | | |
| QY | 1000 | ----- | 999 | | | | |
| Db | 3072 | CTACAGCTGAAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTTCCAGCA | 3131 | | | | |
| QY | 1000 | ----- | 999 | | | | |
| Db | 3132 | GTTCAGAAAGCAGAAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCT | 3191 | | | | |
| QY | 1000 | ----- | 999 | | | | |
| Db | 3192 | GTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGGAAAGGA | 3251 | | | | |
| QY | 1000 | ----- | 999 | | | | |
| Db | 3252 | CTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCAGAAATGTC | 3311 | | | | |
| QY | 1000 | ----- | 999 | | | | |
| Db | 3312 | ACTCGGCTTCTACGAAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGAACCTG | 3371 | | | | |
| QY | 1000 | -----AGACCCTTGAAAAGACTCCAGGAACTTCAA | 1028 | | | | |
| Db | 3372 | CACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAAGACTCCAGGAACTTCAA | 3431 | | | | |
| QY | 1029 | GAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGGATCCTGG | 1088 | | | | |
| Db | 3432 | GAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGGATCCTGG | 3491 | | | | |
| QY | 1089 | CAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCA | 1148 | | | | |
| Db | 3492 | CAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCA | 3551 | | | | |
| QY | 1149 | CTTCGAGGAGAAAAATTGGCCCTCTGAAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGC | 1208 | | | | |
| Db | 3552 | CTTCGAGGAGAAAAATTGGCCCTCTGAAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGC | 3611 | | | | |
| QY | 1209 | CAGCTTACCACCTTTGGGCATTCAAGTCTCACCGTATAACCTCAGCACTCTGGAAGACCTG | 1268 | | | | |
| Db | 3612 | CAGCTTACCACCTTTGGGCATTCAAGTCTCACCGTATAACCTCAGCACTCTGGAAGACCTG | 3671 | | | | |
| QY | 1269 | AACACCAGATGGAAGCTTCTGCAAGTGGCCGTGAGGACCGAGTCAGGCAGCTGCTGATGAA | 1328 | | | | |
| Db | 3672 | AACACCAGATGGAAGCTTCTGCAAGTGGCCGTGAGGACCGAGTCAGGCAGCTGCTGATGAA | 3731 | | | | |
| QY | 1329 | GCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTTCTTTCCACGTCTGTCCAGGGTCCC | 1388 | | | | |
| Db | 3732 | GCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTTCTTTCCACGTCTGTCCAGGGTCCC | 3791 | | | | |
| QY | 1389 | TGGGAGAGAGCCCATCTCGCCAAAACAAAAGTGCCCTACTATATATCAACACGAGACTCAAA | 1448 | | | | |
| Db | 3792 | TGGGAGAGAGCCCATCTCGCCAAAACAAAAGTGCCCTACTATATCAACACGAGACTCAAA | 3851 | | | | |
| QY | 1449 | ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAAT | 1508 | | | | |
| Db | 3852 | ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAAT | 3911 | | | | |
| QY | 1509 | GTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTTGC | 1568 | | | | |
| Db | 3912 | GTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTTGC | 3971 | | | | |
| QY | 1569 | TTGGATCTCTTGAGCCCTGTTCAGCTGCATGTGATGCCCTTGGACGACACACCTCAAGCAA | 1628 | | | | |
| Db | 3972 | TTGGATCTCTTGAGCCCTGTTCAGCTGCATGTGATGCCCTTGGACGACACACCTCAAGCAA | 4031 | | | | |
| QY | 1629 | AATGACCAAGCCCATGGATATCCTGACAGATTATTAATTGTTTGACCACCTATTATGACCGC | 1688 | | | | |
| Db | 4032 | AATGACCAAGCCCATGGATATCCTGACAGATTATTAATTGTTTGACCACCTATTATGACCGC | 4091 | | | | |
| QY | 1689 | CTGGAGCAAGAGCAACAATAATTTGGTCAACAGTCCCTCTCTCGTGGATATGTGTCTGAAC | 1748 | | | | |

Db 4092 CTGGAGCAAGAGCACAAATTGGTCAACGTCCCTCTCTCGTGGATATGTGTCTGAAC 4151
QY 1749 TGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGTCTCTCTTTTAAA 1808
Db 4152 TGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGTCTCTCTTTAAA 4211
QY 1809 ACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGATACCTTTTCAAG 1868
Db 4212 ACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGATACCTTTTCAAG 4271
QY 1869 CAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGCCCTCCTTCTGCATGAT 1928
Db 4272 CAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGCCCTCCTTCTGCATGAT 4331
QY 1929 TCTATCCCAAATCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGCAGTAACATTGAG 1988
Db 4332 TCTATCCCAAATCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGCAGTAACATTGAG 4391
QY 1989 CCAAGTGTCCGGA 2001
Db 4392 CCAAGTGTCCGGA 4404

RESULT 12
ABK81998
ID ABK81998 standard; DNA; 5339 BP.
XX
AC ABK81998;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding mini-dystrophin protein deltaR2-R21.
XX
KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
DS Homo sapiens.
DS Synthetic.
XX
FN WO200229056-A2.
XX
PN 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US311126.
XX
PR 06-OCT-2000; 2000US-238848P.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Chamberlain JS, Harper SQ;
XX
DR WPI; 2002-435334/46.
XX

PT A composition for preparing therapeutic drugs, has a mini-dystrophin
PT peptide comprising a specific number of spectrin-like repeat domains,
PT or a nucleic acid sequence encoding the mini-dystrophin peptide -
XX
PS Example 6; Fig 13; 145pp; English.
XX
CC The invention describes a composition comprising a mini-dystrophin
CC peptide comprising a spectrin-like repeat domain, where the domain
CC comprises n spectrin-like repeats, and contains no more than n
CC spectrin-like repeats, where n is an even number between 4-24, or a
CC nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin
CC peptide or the polynucleotide encoding it is useful as a medicament,
CC for preparing a drug for therapeutic application and in the preparation
CC of a composition for treatment of muscle disease, e.g. Duchenne's
CC muscular dystrophy (DMD). This sequence represents a mini-dystrophin
CC sequence of the invention.
XX
SQ Sequence 5339 BP; 1638 A; 1191 C; 1187 G; 1323 T; 0 other;

Query Match 62.7%; Score 1254; DB 24; Length 5339;
Best Local Similarity 77.9%; Pred. No. 0;
Matches 1621; Conservative 0; Mismatches 380; Indels 79; Gaps 6;
QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAA 60
Db 1199 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAA 1258
QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCAACAAGGAGAGATTTCCTAAT 120
Db 1259 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCAACAAGGAGAGATTTCCTAAT 1318
QY 121 GATGTGGAAGTGGTGAAGACCAAGTTCATACTCATGAGGGGTACATGATGGATTGACA 180
Db 1319 GATGTGGAAGTGGTGAAGACCAAGTTCATACTCATGAGGGGTACATGATGGATTGACA 1378
QY 181 GCCCATCAGGGCCGGTTGGTAAATATTCTACAAATTGGGAAGTAAAGCTGATTGGAACAGGA 240
Db 1379 GCCCATCAGGGCCGGTTGGTAAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGA 1438
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 300
Db 1439 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 1498
QY 301 TGGGAATGCTCAGGGTAGCTAGCATGGAACCAAAAGCAATTTACATAGAGTTTAAATG 360
Db 1499 TGGGAATGCTCAGGGTAGCTAGCATGGAACCAAAAGCAATTTACATCATAGATTACTG 1558
QY 361 GATCTCCAGAATC-GAAACTGAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAAGA 419
Db 1559 CAACAGTTCCTCCCTGGACCTGGAAAAGTTTCTTGCTGCTGTACAGAAGCTGAAACAAC 1618
QY 420 ACAAGGAAATGGAGGAGGCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTA 479
Db 1619 GCCAATGTCTACAGGATGCTACCCGTAAGGAAAGGCTCTAGAAGACTCCAAGSGAGTA 1678
QY 480 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAAACAAGTCAGGGTCAATTTCTC 539
Db 1679 AAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAATTGAAGCTCACACAGATGTT 1738
QY 540 ACTCACATGGTGGTGTAG-----TTGATGAATCTAGTGGAGATCACGCAACT 587
Db 1739 TATCACAACTGGATGAAACAGCCAAAATACTCTGATCCCTGGAAAGTTTCCGATGAT 1798
QY 588 GCTGCTTTTGAAGAACAACTTAAGG---TATTGGGAGATCGATGGGCAACATCTGTAGA 644
Db 1799 GCAGTCTGTTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAGTGAACCTTCGAAA 1858
QY 645 TGGACAGAAGACCGCTGGGTTCTTTTACAAGACATCTTCTCAAATGGCAACGTTTACT 704
Db 1859 AAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCAC 1918
QY 705 GAAGAACAGTGCCTTTTATGTGATGGCTTTTCAGAAAAGAAAGATGCAGTGAACAAGATT 764
Db 1919 CTTTCTCTGCAGGAACCTTCTGCTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAG 1978
QY 765 CACACAACCTGGCTTTAAAGATCAAAAATGAAATGTTATCAAGTCTTCAAAAACCTGCCGTT 824
Db 1979 GCACCTATTGGAGGCGACTTTCCAGCAGTTCCAGAAGCAGAACGATGTACATAGGGCCTC 2038
QY 825 TTAAAGCGGATCTAGAAAAGAAAAGCAATCCATGGGCAAACTGTA----- 871
Db 2039 AAGAGGGAATTGAAAACCTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATA 2098
QY 872 -TTCACTCAAACAAGATCTTCTTCAACACTGAAGATAAAGTCAAGTACCCAGAACGCG 930
Db 2099 TTTCTGACAGAGCAGCCTTTTGAAGGACTAGAGAAACTTACCAGGAGCCCAAGAGAGCTG 2158
QY 931 AAGCATGGCTGGA-----TAACCTTGGCCCGGTGTTGGG 963
Db 2159 CCTCCTGAGGAGAGAGCCCGAAGTGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTC 2218
QY 964 ATAATTAGTCCAAAAAAGTTGAA-----AAGAGTACAGCACAG 1001

Db 2219 AATACTGAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAG 2278
QY 1002 ACCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGC 1061
Db 2279 ACCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGC 2338
QY 1062 CAAGCTGAGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTC 1121
Db 2339 CAAGCTGAGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTC 2398
QY 1122 CAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAAC 1181
Db 2399 CAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAAC 2458
QY 1182 GTGAGCCACGTCGAATGACCTTGCTGCGCAGCTTACCACCTTTGGGCATTGACTCTCACCG 1241
Db 2459 GTGAGCCACGTCGAATGACCTTGCTGCGCAGCTTACCACCTTTGGGCATTGACTCTCACCG 2518
QY 1242 TATAACCTCAGCACTCTGGAAGACCTGGAACACCAAGGACTTTGGTCCAGCATCTCAGCAC 1301
Db 2519 TATAACCTCAGCACTCTGGAAGACCTGGAACACCAAGGACTTTGGTCCAGCATCTCAGCAC 2578
QY 1302 GAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGCACTTTGGTCCAGCATCTCAGCAC 1361
Db 2579 GAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGCACTTTGGTCCAGCATCTCAGCAC 2638
QY 1362 TTTCTTTCCACGTCCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCC 1421
Db 2639 TTTCTTTCCACGTCCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCC 2698
QY 1422 TACTATATCAACACGAGACTCAAAACRACTTGCTGGGACCATCCCAAAATGACAGAGCTC 1481
Db 2699 TACTATATCAACACGAGACTCAAAACRACTTGCTGGGACCATCCCAAAATGACAGAGCTC 2758
QY 1482 TACCAGTCTTTAGCTGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAA 1541
Db 2759 TACCAGTCTTTAGCTGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAA 2818
QY 1542 CTCCGAAGACTGCGAAGAGGCCCTTTGCTGGATCTCTTGAGCCCTGTCAGCTGCGATGAT 1601
Db 2819 CTCCGAAGACTGCGAAGAGGCCCTTTGCTGGATCTCTTGAGCCCTGTCAGCTGCGATGAT 2878
QY 1602 GCCTTGGACCAAGCAACCTCAAGCAAAATGACCCCATGGATATCCTGCAGATTATT 1661
Db 2879 GCCTTGGACCAAGCAACCTCAAGCAAAATGACCCCATGGATATCCTGCAGATTATT 2938
QY 1662 AATTGTTTGACCACTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTC 1721
Db 2939 AATTGTTTGACCACTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTC 2998
QY 1722 CCTCTCTGCGTGGATATGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACA 1781
Db 2999 CCTCTCTGCGTGGATATGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACA 3058
QY 1782 GGGAGGATCCGTGCTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTG 1841
Db 3059 GGGAGGATCCGTGCTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTG 3118
QY 1842 GAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCCAG 1901
Db 3119 GAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCCAG 3178
QY 1902 CGCAGGCTGGGCTCTCTGATGATTCTATCCAAATTCGAAGACAGTGGGTGAAGTT 1961
Db 3179 CGCAGGCTGGGCTCTCTGATGATTCTATCCAAATTCGAAGACAGTGGGTGAAGTT 3238
QY 1962 GCATCCTTTGGGGCAGTAACATTGAGCCCAAGTGTCCGGA 2001
Db 3239 GCATCCTTTGGGGCAGTAACATTGAGCCCAAGTGTCCGGA 3278

RESULT 13

AAD37238
ID AAD37238 standard; DNA; 3531 BP.
XX
AC AAD37238;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta3531.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW ageno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
WP I; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 50-51; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3531 containing nucleotides 1-1341 (N-terminus,
CC hinge H1 and rod R1), 8059-10227 (rods R22, R23 and R24, hinge H4 and
CC CR domain) and 11047-11058 (dystrophin last 3 amino acids).
SQ Sequence 3531 BP; 1071 A; 809 C; 824 G; 827 T; 0 other;
Query Match 62.5%; Score 1251.2; DB 24; Length 3531;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1623; Conservative 0; Mismatches 378; Indels 85; Gaps 6;
QY 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA 60
Db 1000 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA 1059
QY 61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGCAACAAGGAGAGATTCTTAAT 120
Db 1060 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGCAACAAGGAGAGATTCTTAAT 1119
QY 121 GATGTGGAAGTGGTGAAGACCAGTTTCATCTACTCATGAGGGGTACATGATGGATTGACA 180
Db 1120 GATGTGGAAGTGGTGAAGACCAGTTTCATCTACTCATGAGGGGTACATGATGGATTGACA 1179
QY 181 GCCCATCAGGGCCGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240
Db 1180 GCCCATCAGGGCCGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 1239
QY 241 AAATTATCAGAAGATGAAGAAACTGAAGTACAGAGCAGATGATCTCCTAAATTCAAGA 300

Db 1240 AAATTATCAGAAGATGAAGAACTGAAAGTACAAGAGCAGATGAATCTCCTAAATTCAGA 1299
QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATG 360
Db 1300 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAACTCATAGA 1359
QY 361 GATCTCCAGAA-----TCGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAA 413
Db 1360 TTACTGCAACAGTTCCCTCGACCTGGAAAAGTTTCTTGCTGGCTTACAGAAAGTCAA 1419
QY 414 GAAAGAACAGGAAAAATGGAGGAAGACCTCTTGGACCTTGATCTTTGAAGACCTTAAAAAGC 473
Db 1420 ACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAG 1479
QY 474 CAAGTACAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAGGTCAT 533
Db 1480 GGAGTAAAAAGACTGATGAACAATGGCAAGACCTCCAAGTGAAAATTTGAAGCTCACACA 1539
QY 534 TCTCTCACTCACATGGTGGTGTAGTTGATGA-----ATCTAGTGGAGATCAC 581
Db 1540 GATGTTTTATCACAACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCC 1599
QY 582 GCAACTGCTGCTTTGGAAGAACAACTTAAGG--TATTGGAGATCGATGGGCACAAATC 638
Db 1600 GATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAGTGAACCT 1659
QY 639 TGTAGATGSAAGAACAGCCGCTGGGTTCTTTTACAAGACATCCTTCTCAAATGGCAACGT 698
Db 1660 CGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAAGCT 1719
QY 699 CTTACTGAGAACACAGTGCCTTTTTTAGTGCATGGCTTTCAGAAAAAGAAAGATGCAGTGAAC 758
Db 1720 CTGCACCTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 1779
QY 759 AAGATTACACAACTGGCTTTAAAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACTG 818
Db 1780 CGGCAGGCACCTATTGGAGGGACCTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGG 1839
QY 819 GCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAGCAATCCATGGGCAAACTGTA----- 871
Db 1840 GCCTTCAAGAGGGAATTGAAAACTTAAAGAACCTGTATCATGAGTACTCTTTGAGACTGTA 1899
QY 872 -----TTCACTCAAAACAAGATCTTCTTTCAACA CTGAAGAAATAAGTCAAGTACCAG 924
Db 1900 CGAATATTTCTGACAGAGCAGCCTTTTGGAAAGGACTAGAGAACTCTACAGGAGCCACAG 1959
QY 925 AGACGGAAGCATGGCTGGA-----TAACTTTGCCCGGT 957
Db 1960 GAGCTGCCTCCTGAGGAGAGAGCCCGAAGTGTCACTCGGCTTCTACGAAAGCAGGCTGAG 2019
QY 958 GTTGGGATAATTTAGTCCAAAAACTTGAA-----AAGAGTACA 995
Db 2020 GAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 2079
QY 996 GCACAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAG 1055
Db 2080 GATGAGACCCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAG 2139
QY 1056 CTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGAC 1115
Db 2140 CTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGAC 2199
QY 1116 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAA 1175
Db 2200 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAA 2259
QY 1176 GAGAACGTGAGCCACGTCAATGACCTTGTGCGCCAGCTTACCAGTTTGGGCAATTCAGCTC 1235
Db 2260 GAGAACGTGAGCCACGTCAATGACCTTGTGCGCCAGCTTACCAGTTTGGGCAATTCAGCTC 2319
QY 1236 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTG 1295

Db 2320 TCACCGTATAAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTG 2379
QY 1296 GCCGTCGAGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGACTTTGGTCCAGCATCT 1355
Db 2380 GCCGTCGAGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGACTTTGGTCCAGCATCT 2439
QY 1356 CAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 1415
Db 2440 CAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 2499
QY 1416 GTGCCCTACTATATCAACCACGAGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACA 1475
Db 2500 GTGCCCTACTATATCAACCACGAGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACA 2559
QY 1476 GAGCTCTACAGTCTTTAGCTGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCC 1535
Db 2560 GAGCTCTACAGTCTTTAGCTGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCC 2619
QY 1536 ATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTTCAGCTGCA 1595
Db 2620 ATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTTCAGCTGCA 2679
QY 1596 TGTGATGCCTTGGACCAGCACAACTCAAGCAAAAATGACCAGCCCATGGATATCCTGCGAG 1655
Db 2680 TGTGATGCCTTGGACCAGCACAACTCAAGCAAAAATGACCAGCCCATGGATATCCTGCGAG 2739
QY 1656 ATTATTAAATTGTTTGACCACCTATTATGACCCGCTGGAGCAAGAGCACAACTTTGGTC 1715
Db 2740 ATTATTAAATTGTTTGACCACCTATTATGACCCGCTGGAGCAAGAGCACAACTTTGGTC 2799
QY 1716 AACGTCCCTCTCTGCGTGGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGA 1775
Db 2800 AACGTCCCTCTCTGCGTGGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGA 2859
QY 1776 CGAACAGGAGGATCCGTGTCCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 1835
Db 2860 CGAACAGGAGGATCCGTGTCCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 2919
QY 1836 CATTTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGT 1895
Db 2920 CATTTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGT 2979
QY 1896 GACCAGCGCAGGCTGGGCTCTCTCTGATGATTTCTATCCAAATTCGAAGACAGTTGGGT 1955
Db 2980 GACCAGCGCAGGCTGGGCTCTCTCTGATGATTTCTATCCAAATTCGAAGACAGTTGGGT 3039
QY 1956 GAAGTTGCATCCTTTGGGGGCAGTAACATTGAGCCAAAGTGTCCGGA 2001
Db 3040 GAAGTTGCATCCTTTGGGGGCAGTAACATTGAGCCAAAGTGTCCGGA 3085

RESULT 14
AAD37258
ID AAD37258 standard; DNA; 4498 BP.
XX
AC AAD37258;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-3531.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Unidentified.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.

XX
PR 28-APR-2000; 2000US-200777P.

PA (XIAO/) XIAO X.

PI Xiao X;

DR WPI; 2002-049342/06.

PT New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,

PT rod repeats, H1 and H4 domains and a cysteine rich domain of a

PT dystrophin gene -
....

PS Example 1; Page 62-63; 71pp; English.

The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is AAV vector plasmid construct containing human dystrophin minigenes, a muscle creatine kinase (MCK) promoter and a small polyA signal sequence.

SQ Sequence 4498 BP; 1251 A; 1118 C; 1123 G; 1006 T; 0 other;

Query Match 62.5%; Score 1251.2; DB 24; Length 4498;

Best Local Similarity 77.8%; Pred. No. 0;

Matches 1623; Conservative 0; Mismatches 378; Indels 85; Gaps 6;

Qy 1 GGCAGTTCA TTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGAA 60

Db 1757 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA 1816

QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACCAAGGAGAGATTTCTTAAT 120

Db 1817 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTCTAAT 1876

QY 121 GATGTGGAAGTGGTGAAAGACCAGTTTCATACTCATGAGGGGTACATGATGGATTGACA 180

Db 1877 GATGTGGAAGTCGTGAAAGACCAGTTTCATCTCATGAGGGTACATGATGGATTGACA 1936

QY 181 GCCCATCAGGCCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240

Db 1937 GCCCATCAGGCCGGTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 1996

QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAACTCTCCTAAATTCAAGA 300

Db 1997 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATCAAGA 2056

QY 301 TGGGAATGCCCTCAGGGTAGCTAGCATGGAATAAACAAGCAATTTACATAGAGTTTAAATG 360

Db 2057 TGGGAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAACTCATAGA 2116

QY 361 GATCTCCAGAA-----TCGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAA 413

D**b** 2117 T T A C T G C A A C A G T T C C C C C T G G A C C T G G A A A G T T T C T T G C C T G G C T T A C A G A A G C T G A A 21176

QY 414 GAAAGAACAGGAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTGAAGACCTAAACGC 473

Db 2177 ACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAAGACTCCAAG 2236

474 CAAGTACAACAACATAAGGTGCTTCAAGGAAGATCTAGAACAAAGAACAGTCAGGGTCAAT 533

Db 2237 GGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACA 2296

Qy 534 TCTCTCACTCACATGGTGGTGGTAGTTGATGA-----ATCTAGTGGAGATCAC 581

| | | | |
|----|------|---|------|
| Db | 2297 | GATGTTTATCACAACCTGGATGAAACACAGCCAAAAAATCCTGAGATCCCTGGAAGTTCC | 2355 |
| QY | 582 | GCAACTGCTGCTTTGGAAGAACAACTTAAAG--TATTGGGAGATCGATGGGCAAAACATC | 638 |
| Db | 2357 | GATGATGCAGTCTCTTACAAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTT | 2416 |
| QY | 639 | TGTAGATGGACAGAGAAGACCGCTGGGTTCTTTTAGTGCATGGCTTTCAGAAAAAGAGATGCACTGAAC | 698 |
| Db | 2417 | CGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCCAGTGAAGCGT | 2476 |
| QY | 699 | CTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGAAAAAGAGATGCACTGAAC | 758 |
| Db | 2477 | CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC | 2536 |
| QY | 759 | AAGATTACACAACACTGGCTTTAAAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACTG | 818 |
| Db | 2537 | CGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGG | 2596 |
| QY | 819 | GCCGTTTTTAAAGCGGATCTAGAAAAAGHAAAAAGCAATCCATGGGCAAACTGTA----- | 871 |
| Db | 2597 | GCCTTTCAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTTGAGACTGTA | 2656 |
| QY | 872 | -----TTCACTCAAACAAGATCTTTCTTTCAACACTGAAGAAATAAGTCACTGACCCAGA | 924 |
| Db | 2657 | CGAATATTTCTGACAGAGCAGCCTTTTGGAAAGGACTAGAGAAACTCTACAGGAGCCCGA | 2716 |
| QY | 925 | AGACGGAAGCATGGCTGGA-----TAACTTTGCCCGGT | 957 |
| Db | 2717 | GAGCTGCCCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG | 2776 |
| QY | 958 | GTGGGATAATTAGTCCAAAAACTTGAA-----AAGAGTACA | 995 |
| Db | 2777 | GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATA | 2836 |
| QY | 996 | GCACAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAG | 1055 |
| Db | 2837 | GATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAG | 2896 |
| QY | 1056 | CTGCGCCAAGCTGAGGTGATCAAGGGATCTTGGCAGCCCGTGGCGATCTCTCATTTGAC | 1115 |
| Db | 2897 | CTGCGCCAAGCTGAGGTGATCAAGGGATCTTGGCAGCCCGTGGCGATCTCTCATTTGAC | 2956 |
| QY | 1116 | TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTCGAGGAGAAATTCGCTCTGAAA | 1175 |
| Db | 2957 | TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTCGAGGAGAAATTCGCTCTGAAA | 3016 |
| QY | 1176 | GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGCATTTCAGCTC | 1235 |
| Db | 3017 | GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGCATTTCAGCTC | 3076 |
| QY | 1236 | TCACCGTATAACCTCAGCACCTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTG | 1295 |
| Db | 3077 | TCACCGTATAACCTCAGCACCTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTG | 3136 |
| QY | 1296 | GCCGTCGAGGACCGAGTCAGGCAGCTGCAATGAAGCCCAAGGACCTTTGGTCCAGCATCT | 1355 |
| Db | 3137 | GCCGTCGAGGACCGAGTCAGGCAGCTGCAATGAAGCCCAAGGACCTTTGGTCCAGCATCT | 3196 |
| QY | 1356 | CAGCACCTTCTTTCCACGTCTGTCCAGGTCCTGCGGAGAGAGCCATCTCGCCAAACAAA | 1415 |
| Db | 3197 | CAGCACCTTCTTTCCACGTCTGTCCAGGTCCTGCGGAGAGAGCCATCTCGCCAAACAAA | 3256 |
| QY | 1416 | GTGCCCTACTATATCAACCAAGAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACA | 1475 |
| Db | 3257 | GTGCCCTACTATATCAACCAAGAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACA | 3316 |
| QY | 1476 | GAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCC | 1535 |
| Db | 3317 | GAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCC | 3376 |
| QY | 1536 | ATGAAACTCCGAAGACTGCAGAAGGCCCTTTTGCTTGGATCTCTTGAGCCTGTTCAGCTGCA | 1595 |
| Db | 3377 | ATGAAACTCCGAAGACTGCAGAAGGCCCTTTTGCTTGGATCTCTTGAGCCTGTTCAGCTGCA | 3436 |

QY 1596 TGTGATGCCCTTGACCAGCACAACTCAAGCAAAATGACCAGCCCATGGATATCCTGCGAG 1655
DB TGTGATGCCCTTGACCAGCACAACTCAAGCAAAATGACCAGCCCATGGATATCCTGCGAG 3496
QY 1656 ATTATTAATTGTTGACCACACTATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTC 1715
DB ATTATTAATTGTTGACCACACTATTATGACCGCCTGGAGCAAGAGCACAAATTTGGTC 3556
QY 1716 AACGTCCCTCTCTGCGTGGATATGTCGTGAACCTGGCTGCTGTAATGTTTATGATACGGGA 1775
DB AACGTCCCTCTCTGCGTGGATATGTCGTGAACCTGGCTGCTGTAATGTTTATGATACGGGA 3616
QY 1776 CGAACAGGGAGATCCGTGTCCTGTCCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 1835
DB CGAACAGGGAGATCCGTGTCCTGTCCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 3676
QY 1836 CATTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGT 1895
DB CATTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGT 3736
QY 1896 GACCAGCGAGGTGGGCTCCTTCTGCGATGATTCATCCAAATCCAAAGACAGTTGGGT 1955
DB GACCAGCGAGGTGGGCTCCTTCTGCGATGATTCATCCAAATCCAAAGACAGTTGGGT 3796
QY 1956 GAAGTTGCATCCTTTGGGGCGAGTAACATTTGAGCCCAAGTGTCGGGA 2001
DB GAAGTTGCATCCTTTGGGGCGAGTAACATTTGAGCCCAAGTGTCGGGA 3842

RESULT 15
AAD37230
ID AAD37230 standard; DNA; 4182 BP.
XX
AC AAD37230;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta4173.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 43-44; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control

CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta4173 containing nucleotides 1-1992 (N-terminus,
CC hinge H1 and rods R1, R2 and R3), 8059-10227 (rods R22, R23 and R24,
CC hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids).
XX
SQ Sequence 4182 BP; 1309 A; 927 C; 970 G; 976 T; 0 other;
Query Match 62.2%; Score 1245; DB 24; Length 4182;
Best Local Similarity 73.1%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 736; Gaps 2;
QY 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA 60
DB 1000 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA 1059
QY 61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAAT 120
DB 1060 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAAT 1119
QY 121 GATGTGGAAGTGGTGAAGACCAGTTTCATCTACTCATGAGGGGTACATGATGATTGACA 180
DB 1120 GATGTGGAAGTGGTGAAGACCAGTTTCATCTACTCATGAGGGGTACATGATGATTGACA 1179
QY 181 GCCCATCAGGCGCGGTTGGTAATATTTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240
DB 1180 GCCCATCAGGCGCGGTTGGTAATATTTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 1239
QY 241 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGA 300
DB 1240 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGA 1299
QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTACATAGATTTTAATG 360
DB 1300 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTACATAGATTTTAATG 1359
QY 361 GATCTCCAGATC-GAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGA 419
DB 1360 GATCTCCAGATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGA 1419
QY 420 ACAAGGAAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTA 479
DB 1420 ACAAGGAAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTA 1479
QY 480 CAACAACATAAGGTGCTTCAAGAGAGATCTAGAACAAAGAACAACTCAGGGTCAATTCCTC 539
DB 1480 CAACAACATAAGGTGCTTCAAGAGAGATCTAGAACAAAGAACAACTCAGGGTCAATTCCTC 1539
QY 540 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAGCAAACTGCTGCTTTGGAA 599
DB 1540 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAGCAAACTGCTGCTTTGGAA 1599
QY 600 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACACATCTGTAGATGGACAGAAGACCGC 659
DB 1600 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACACATCTGTAGATGGACAGAAGACCGC 1659
QY 660 TGGGTTCTTTTACAAGACATCCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTT 719
DB 1660 TGGGTTCTTTTACAAGACATCCTTCTCAAATGGCAACGCTCTTACTGAAGAACAGTGCCTT 1719
QY 720 TTTAGTGCATGGCTTTTCAGAAAAAGAGATGCAGTGAACAAGATTCACACAACACTGGCTTT 779
DB 1720 TTTAGTGCATGGCTTTTCAGAAAAAGAGATGCAGTGAACAAGATTCACACAACACTGGCTTT 1779
QY 780 AAAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTA 839
DB 1780 AAAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTA 1839
QY 840 GAAAAGAAAAAGCAATCCATGGGCAAACTGTATTCTACTCAAAAGAGATCTTCTTTCAACA 899
DB 1840 GAAAAGAAAAAGCAATCCATGGGCAAACTGTATTCTACTCAAAAGAGATCTTCTTTCAACA 1899

QY 900 CTGAAGAAATAAGTCAGTGACCCAGAAAGACGGAAAGCATGGCTGGATAAACTTTGCCCCGGTGT 959
Db |||||
1900 CTGAAGAAATAAGTCAGTGACCCAGAAAGACGGAAAGCATGGCTGGATAAACTTTGCCCCGGTGT 1959
QY 960 TGGGATAATTTAGTCCAAAACCTTGAAAAGAGTACAGCAC----- 999
Db |||||
1960 TGGGATAATTTAGTCCAAAACCTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAA 2019
QY 1000 ----- 999
Db |||||
2020 CAGTTCGCCCTGGACCTGGAAAAGTTTCTTGCTGGCTTACAGAAAGCTGAACAACACTGCC 2079
QY 1000 ----- 999
Db |||||
2080 AATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCTPAGAAGACTCCAAGGGAGTAAAA 2139
QY 1000 ----- 999
Db |||||
2140 GAGCTGATGAACAATAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTAT 2199
QY 1000 ----- 999
Db |||||
2200 CACAACCTGGATGAACAACAGCCAAAAAATCCTGAGATCCCTGGAAGTTCCCGATGATGCA 2259
QY 1000 ----- 999
Db |||||
2260 GTCCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAG 2319
QY 1000 ----- 999
Db |||||
2320 TCTCTCAACATTAGGTCCCATTTTGAAGCCAGTTCTGACCAGTGAAGCGTCTGACACCTT 2379
QY 1000 ----- 999
Db |||||
2380 TCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAAGATGATGAATTAAGCCGCGCAGGCA 2439
QY 1000 ----- 999
Db |||||
2440 CCTATTGGAGCGGACTTTCAGCAGCTTCAAGAGCAGAACGATGTACATAGGGCCCTTCAAG 2499
QY 1000 ----- 999
Db |||||
2500 AGGGAATTGAAAACATAAGAAACCTGTAATCATGAGTACTCTTTGAGACTGTACGAAATATTT 2559
QY 1000 ----- 999
Db |||||
2560 CTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAAACCTCTACCAGAGCCCAGAGAGCTGCGCT 2619
QY 1000 ----- 999
Db |||||
2620 CCTGAGGAGAGAGCCCAAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAAT 2679
QY 1000 -----AGACC 1004
Db |||||
2680 ACTGAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAAATAGATGAGACC 2739
QY 1005 CTTGAAAGACTCCAGGAACCTTCAAGAGGCGCACGGATGAGCTGGACCTCAAGCTGCGCCAA 1064
Db |||||
2740 CTTGAAAGACTCCAGGAACCTTCAAGAGGCGCACGGATGAGCTGGACCTCAAGCTGCGCCAA 2799
QY 1065 GCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGGATCTCCTCATTTGACTCTCTCCAA 1124
Db |||||
2800 GCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGGATCTCCTCATTTGACTCTCTCCAA 2859
QY 1125 GATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGGCTCTGAAAGAGAACCTG 1184
Db |||||
2860 GATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGGCTCTGAAAGAGAACCTG 2919
QY 1185 AGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCCGTAT 1244
Db |||||
2920 AGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCCGTAT 2979

QY 1245 AACCTCAGCACTCTGGAGACCTGGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAG 1304
Db |||||
2980 AACCTCAGCACTCTGGAGACCTGGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAG 3039
QY 1305 GACCAGTTCAGGCAGCTGCAATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTT 1364
Db |||||
3040 GACCAGTTCAGGCAGCTGCAATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTT 3099
QY 1365 CTTTCCACGTCGTGCCAGGTCCCTGGGAGAGAGCCATCTGCCAAAACAAAGTGCCCTAC 1424
Db |||||
3100 CTTTCCACGTCGTGCCAGGTCCCTGGGAGAGAGCCATCTGCCAAAACAAAGTGCCCTAC 3159
QY 1425 TATATCAACCACGAGACTCAAAACAACTTGTCTGGACCATCCAAAATGACAGAGCTCTAC 1484
Db |||||
3160 TATATCAACCACGAGACTCAAAACAACTTGTCTGGACCATCCAAAATGACAGAGCTCTAC 3219
QY 1485 CAGTCTTTAGCTGACCTGAATAATGTCAAGTTCTCAGCTTATAGGACTGCCATGAAACTC 1544
Db |||||
3220 CAGTCTTTAGCTGACCTGAATAATGTCAAGTTCTCAGCTTATAGGACTGCCATGAAACTC 3279
QY 1545 CGAAGACTGCAGAAGGCCCTTTGCTTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCC 1604
Db |||||
3280 CGAAGACTGCAGAAGGCCCTTTGCTTTGGATCTCTTGAGCCCTGTGAGCTGCATGTGATGCC 3339
QY 1605 TTGGACCAGCACAACTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAAT 1664
Db |||||
3340 TTGGACCAGCACAACTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAAT 3399
QY 1665 TGTTTGACCACATAATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTT 1724
Db |||||
3400 TGTTTGACCACATAATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTT 3459
QY 1725 CTCTGCGTGGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGG 1784
Db |||||
3460 CTCTGCGTGGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGG 3519
QY 1785 AGGATCCGTGTCCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAA 1844
Db |||||
3520 AGGATCCGTGTCCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAA 3579
QY 1845 GACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCGC 1904
Db |||||
3580 GACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCGCGC 3639
QY 1905 AGGCTGGGCTCCTTCTGCAATGATTTAGCCAAAGTGTCCGGA 2001
Db |||||
3640 AGGCTGGGCTCCTTCTGCAATGATTTAGCCAAAGTGTCCGGA 3736
QY 1965 TCCTTTGGGGCAGTAACATTTAGCCAAAGTGTCCGGA 2001
Db |||||
3700 TCCTTTGGGGCAGTAACATTTAGCCAAAGTGTCCGGA 3736

Search completed: February 1, 2004, 12:05:44
Job time : 515.292 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: February 1, 2004, 11:11:40 ; Search time 123.697 Seconds
(without alignments)
7140.092 Million cell updates/sec

Title: US-09-845-416-14_COPY_1000_3000
Perfect score: 2001
Sequence: 1 ggcagttcattgatggagag.....cattgagccaagtgtccgga 2001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|---------------------|
| 1 | 1002.6 | 50.1 | 5952 | 4 | US-09-687-875A-1 |
| 2 | 991.6 | 49.6 | 13977 | 4 | US-09-484-970B-60 |
| 3 | 869.8 | 43.5 | 19307 | 3 | US-08-836-022A-10 |
| 4 | 869.8 | 43.5 | 19307 | 3 | US-09-427-048A-10 |
| 5 | 490.6 | 24.5 | 6045 | 4 | US-09-091-501B-7 |
| 6 | 490.6 | 24.5 | 10320 | 4 | US-09-091-501B-9 |
| 7 | 79.4 | 4.0 | 200 | 4 | US-09-091-501B-5 |
| 8 | 78.6 | 3.9 | 200 | 4 | US-09-091-501B-4 |
| 9 | 78.6 | 3.9 | 200 | 4 | US-09-091-501B-6 |
| 10 | 74.6 | 3.7 | 7218 | 1 | US-08-232-463-14 |
| 11 | 44.2 | 2.2 | 2574 | 4 | US-09-668-313A-10 |
| 12 | 42.8 | 2.1 | 1179 | 4 | US-09-107-532A-1186 |
| 13 | 42.8 | 2.1 | 1690 | 4 | US-09-620-312D-69 |
| 14 | 42.8 | 2.1 | 7812 | 3 | US-09-368-590-1 |
| 15 | 42.4 | 2.1 | 2704 | 3 | US-08-857-076-44 |
| 16 | 42.4 | 2.1 | 3499 | 3 | US-08-857-076-43 |
| 17 | 39.4 | 2.0 | 2223 | 1 | US-08-257-073-4 |
| 18 | 38.8 | 1.9 | 289 | 3 | US-09-007-005-17 |
| 19 | 38.8 | 1.9 | 289 | 3 | US-09-244-796-17 |
| 20 | 38.8 | 1.9 | 1821 | 4 | US-08-477-831C-1 |
| 21 | 38.8 | 1.9 | 1885 | 4 | US-08-477-831C-9 |
| 22 | 38.8 | 1.9 | 1896 | 4 | US-08-477-831C-10 |
| 23 | 38.8 | 1.9 | 1961 | 4 | US-08-477-831C-8 |
| 24 | 38.8 | 1.9 | 2968 | 4 | US-08-477-831C-13 |
| 25 | 38.8 | 1.9 | 3044 | 4 | US-08-477-831C-12 |
| 26 | 38.4 | 1.9 | 7672 | 4 | US-09-220-132-24 |
| 27 | 38.2 | 1.9 | 428 | 4 | US-09-668-313A-3 |

| | | | | | | |
|----|------|-----|-------|---|--------------------|-------------------|
| 28 | 38.2 | 1.9 | 1848 | 4 | US-09-134-001C-447 | Sequence 447, App |
| 29 | 38.2 | 1.9 | 4439 | 4 | US-09-668-313A-17 | Sequence 17, Appl |
| 30 | 37.4 | 1.9 | 2082 | 3 | US-08-985-335-4 | Sequence 4, Appli |
| 31 | 37.4 | 1.9 | 2082 | 3 | US-09-410-372-4 | Sequence 4, Appli |
| 32 | 37.2 | 1.9 | 2915 | 4 | US-09-336-115C-5 | Sequence 5, Appli |
| 33 | 37.2 | 1.9 | 3902 | 4 | US-08-961-527-212 | Sequence 212, App |
| 34 | 36.8 | 1.8 | 4929 | 4 | US-09-620-312D-674 | Sequence 674, App |
| 35 | 36 | 1.8 | 608 | 3 | US-09-385-982-236 | Sequence 236, App |
| 36 | 36 | 1.8 | 2763 | 1 | US-08-248-466B-2 | Sequence 2, Appli |
| 37 | 35.8 | 1.8 | 633 | 4 | US-09-134-001C-578 | Sequence 578, App |
| 38 | 35.6 | 1.8 | 2447 | 2 | US-09-014-969-14 | Sequence 14, Appl |
| 39 | 35.6 | 1.8 | 4868 | 1 | US-08-139-937-12 | Sequence 12, Appl |
| 40 | 35.6 | 1.8 | 4868 | 5 | PCT-US93-11310-12 | Sequence 12, Appl |
| 41 | 35.6 | 1.8 | 8257 | 4 | US-09-595-684B-30 | Sequence 30, Appl |
| 42 | 35.6 | 1.8 | 8789 | 1 | US-08-328-254-5 | Sequence 5, Appli |
| 43 | 35.6 | 1.8 | 10136 | 1 | US-08-353-700-2 | Sequence 2, Appli |
| 44 | 35.6 | 1.8 | 10136 | 5 | PCT-US95-16216-2 | Sequence 2, Appli |
| 45 | 35.4 | 1.8 | 560 | 4 | US-09-220-132-167 | Sequence 167, App |

ALIGNMENTS

RESULT 1
US-09-687-875A-1
; Sequence 1, Application US/09687875A
; Patent No. 6544786
; GENERAL INFORMATION:
; APPLICANT: Xiao, Xiao
; APPLICANT: Liu, Paul
; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPLICED PE
; FILE REFERENCE: 00792
; CURRENT APPLICATION NUMBER: US/09/687,875A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/158,868
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 5952
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2897)..(2898)
; OTHER INFORMATION: S4 junction site
; NAME/KEY: misc feature
; LOCATION: (3198)..(3199)
; OTHER INFORMATION: S2 junction site
US-09-687-875A-1

Query Match 50.1%; Score 1002.6; DB 4; Length 5952;
Best Local Similarity 99.6%; Pred. No. 4.2e-301;
Matches 1005; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

| | | | |
|----|------|--|------|
| QY | 993 | ACAGCACAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGCCCGGATGAGCTGGACCTC | 1052 |
| DB | 3679 | ATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGCCCGGATGAGCTGGACCTC | 3738 |
| QY | 1053 | AAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGAGCCCGTGGCGGATCTCCTCATT | 1112 |
| DB | 3739 | AAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGAGCCCGTGGCGGATCTCCTCATT | 3798 |
| QY | 1113 | GACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCATTCGAGGAGAAATTCGCCCTCTG | 1172 |
| DB | 3799 | GACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCATTCGAGGAGAAATTCGCCCTCTG | 3858 |
| QY | 1173 | AAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAG | 1232 |
| DB | 3859 | AAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAG | 3918 |
| QY | 1233 | CTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGGAACACCATGGAAGCTTCTGCGAG | 1292 |

| | | | |
|----|------|--|------|
| Db | 3919 | CTCTACCGTATAAACCTCAGCACTCTTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAG | 3978 |
| Qy | 1293 | GTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGGACCTTTGGTCCAGCA | 1352 |
| Db | 3979 | GTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGGACCTTTGGTCCAGCA | 4038 |
| Qy | 1353 | TCTCAGCACTTTCTTTCCACGTCGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAC | 1412 |
| Db | 4039 | TCTCAGCACTTTCTTTCCACGTCGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAC | 4098 |
| Qy | 1413 | AAAGTGCCCTACTATATCAACCAAGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATG | 1472 |
| Db | 4099 | AAAGTGCCCTACTATATCAACCAAGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATG | 4158 |
| Qy | 1473 | ACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACT | 1532 |
| Db | 4159 | ACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACT | 4218 |
| Qy | 1533 | GCCATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCCAGCT | 1592 |
| Db | 4219 | GCCATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCCAGCT | 4278 |
| Qy | 1593 | GCATGTGATGCCCTTGGACCAAGCAAACTCAAGCAAAATGACCAAGCCATGGATATCCTG | 1652 |
| Db | 4279 | GCATGTGATGCCCTTGGACCAAGCAAACTCAAGCAAAATGACCAAGCCATGGATATCCTG | 4338 |
| Qy | 1653 | CAGATTATTAATTGTTTGACCACCTATTTATGACCGCCCTGGAGCAAGAGCAACAATTTG | 1712 |
| Db | 4339 | CAGATTATTAATTGTTTGACCACCTATTTATGACCGCCCTGGAGCAAGAGCAACAATTTG | 4398 |
| Qy | 1713 | GTCACGTCCTCTCTGCGTGGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACG | 1772 |
| Db | 4399 | GTCACGTCCTCTCTGCGTGGATATGTGCTGAACCTGGCTGCTGAATGTTTATGATACG | 4458 |
| Qy | 1773 | GGACGAACAGGAGGATCCGTGTCCTGCTCTTTTAAAACTGGCATCATTTCCCTGTGTAAA | 1832 |
| Db | 4459 | GGACGAACAGGAGGATCCGTGTCCTGCTCTTTTAAAACTGGCATCATTTCCCTGTGTAAA | 4518 |
| Qy | 1833 | GCACATTTGGAACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTT | 1892 |
| Db | 4519 | GCACATTTGGAACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTT | 4578 |
| Qy | 1893 | TGTGACCAAGGCTGGGCCTCCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTG | 1952 |
| Db | 4579 | TGTGACCAAGGCTGGGCCTCCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTG | 4638 |
| Qy | 1953 | GGTGAAGTTGCATCCTTTTGGGGGCAGTAACATTGAGCCAAAGTGTCGGGA | 2001 |
| Db | 4639 | GGTGAAGTTGCATCCTTTTGGGGGCAGTAACATTGAGCCAAAGTGTCGGGA | 4687 |

RESULT 2

```

US-09-484-970B-60
; Sequence 60, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmuth, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484, 970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 60
; LENGTH: 13977
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6426186 229357.11CB1
; NAME/KEY: unsure

```

Db 9953 GGGTGAAGTTGCATCCTTTGGGGCAGTAACATTGAGCCAAAGTGTCGGGA 10002

RESULT 3

US-08-836-022A-10/c
; Sequence 10, Application US/08836022A
; Patent No. 6001557
; GENERAL INFORMATION:
; APPLICANT: Trustees of the University of Pennsylvania
; APPLICANT: Wilson, James M.
; APPLICANT: Fisher, Krishna J.
; APPLICANT: Chen, Shu-Jen
; APPLICANT: Weitzman, Matthew
; TITLE OF INVENTION: Improved Adenovirus Virus and
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P O Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,022A

; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/331,381
; FILING DATE: 28-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GNVPN.008PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19307 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
US-08-836-022A-10

Query Match 43.5%; Score 869.8; DB 3; Length 19307;
Best Local Similarity 91.4%; Pred. No. 2e-259;
Matches 922; Conservative 0; Mismatches 87; Indels 0; Gaps 0;
QY 993 ACAGCACAGACCCCTGAAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTC 1052
Db 5701 ATAGATGAAGCTCTTGAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTC 5642
QY 1053 AAGCTGCGCCAAAGCTGAGGTGATCAAGGATCCCTGGCAGCCCGTGGCGGATCTCCTCATT 1112
Db 5641 AAGTTGCGCCAAAGCTGAGGTGATCAAGGATCCCTGGCAGCCCGTGGCGGATCTCCTCATT 5582
QY 1113 GACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGAGAAATTGCGCCTCTG 1172
Db 5581 GACTCTCTGCAAGATCACCTTGAAGAAAGTCAAGGCACCTTCGAGAGAAATTGCACTCTT 5522
QY 1173 AAAGAGAAGCTGAGGCACGTCATATGACCTTGTCTGCCAGCTTACCACTTTGGGCATTGAG 1232
Db 5521 AAAGAGAAGTGAATCGTGTCAATGACCTTGACATCAGCTGACCACTGGGCATTGAG 5462
QY 1233 CTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCGAG 1292

Db 5461 CTCTCACCTTATAACCTCAGCACTTTTGGAAAGATCTGAATACCAGATGGAGGCTTCTACAG 5402
QY 1293 GTGGCCGTGAGGACCGAGTCAAGCAGCTGATGAAGCCCAAGGACTTTGGTCCAGCA 1352
Db 5401 GTGGCTGTGGAGGACCGTGTGAGACAGCTGATGAAGCCCAAGGACTTTGGTCTGCA 5342
QY 1353 TCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCCTTGGGAGAGAGCATTCTGCGCAAAAC 1412
Db 5341 TCCAGCACTTCTCTTTCCACTTCAGTTTCAGGTCCTTGGGAGAGAGCATTCTCAGCAAAAC 5282
QY 1413 AAAGTGCCCTACTATATCAACACGAGACTCAAAACAACTTGTCTGGGACCATCCCAAAATG 1472
Db 5281 AAAGTGCCCTACTATATCAACACGAGACCCCAAAACCACTTGTCTGGGACCATCCCAAAATG 5222
QY 1473 ACAGAGCTCTACCACTCTTTAGCTGACCTGAATAATGTGAGATTCTCAGCTTATAGGACT 1532
Db 5221 ACAGAGCTCTACCACTCTTTAGCTGACCTGAATAATGTGAGATTCTCAGCTTATAGGACT 5162
QY 1533 GCCATGAAACTCCGAAGACTGCAGAGGCCCTTGTCTTGGATCTCTTGGACCTGTGAGCT 1592
Db 5161 GCCATGAAGCTCAGAAAGGCTCCAGAGGCCCTTGTCTTGGATCTCTTGGACCTGTGAGCT 5102
QY 1593 GCATGTGATGCTCTGACCCAGCACCAACCTCAAGCAAAATGACCCAGCCCATGGATATCCTG 1652
Db 5101 GCATGTGATGCTCTGACCCAGCACCAACCTCAAGCAAAATGACCCAGCCCATGGATATCCTG 5042
QY 1653 CAGATTATTAATTGTTGACCACTATTTATGACCGCCTGGAGCAAGAGCACAAATTTG 1712
Db 5041 CAGATAATTAACCTGTTGACTACAATTTATGATCGTCTGGAGCAAGAGCACAAATCTG 4982
QY 1713 GTCAACGTCCTCTCTGCGTGGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACG 1772
Db 4981 GTCAATGTCTCTCTGCTGGATATGTGTCTCAACTGGCTTCTCAATGTTTATGATACG 4922
QY 1773 GGACGAACAGGGAGGATCCGTGCTCTTTTAAACTGGCAATCAATTTCTGTGTAAC 1832
Db 4921 GGACGAACAGGGAGGATCCGTGCTCTTTTAAACTGGCAATCAATTTCTGTGTAAC 4862
QY 1833 GCACATTTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTCAACAGGATTT 1892
Db 4861 GCACATTTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTCAACAGGCTTT 4802
QY 1893 TGTGACCAGCGCAGGCTGGGCTCTCTTCTGATGATTCTATCCAAATCCAAAGACAGTTG 1952
Db 4801 TGTGACCAGCGTAGGCTGGTCTTCTTCTGATGATTCTATTCAATCCCAAGACAGTTG 4742
QY 1953 GGTGAAGTTGCTATCCTTTGGGGGCGAGTAACATTGAGCCGAGTGTCCGGA 2001
Db 4741 GGTGAAGTTGCTATCCTTTGGGGGCGAGTAACATTGAGCCGAGTGTCCGGA 4693

RESULT 4

US-09-427-048A-10/c
; Sequence 10, Application US/09427048A
; Patent No. 6203975
; GENERAL INFORMATION:
; APPLICANT: Trustees of the University of Pennsylvania
; APPLICANT: Wilson, James M.
; APPLICANT: Fisher, Krishna J.
; APPLICANT: Chen, Shu-Jen
; APPLICANT: Weitzman, Matthew
; TITLE OF INVENTION: Improved Adenovirus Virus and
; METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P O Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM: Floppy disk

Db 3843 TGGACGCTGACATGAAGGAGGAGAGTCCGTGCGGAATGGCTGGAAGCCCGTGGGAGACT 3902
QY 1105 TCCTCATTTGACTCTCTCCAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTG 1164
Db 3903 TACTCATTTGACTCGCTGCAGGATCACATTGAAATAATCATGGCATTTAGAGAAGAAATTG 3962
QY 1165 CGCCTCTGAAAGAGAACGCTGAGGCCACGTCATGACCTTGTCTGCCAGCTTACCACTTTGG 1224
Db 3963 CACCAATCAACTTTAAAGTTAAACGGTGAATGATTATCCAGTCAGCTGTCTCCACTTG 4022
QY 1225 GCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTTGAACACCCAGATGGAAGC 1284
Db 4023 ACCTGCATCCCTCTCTAAAGATGTCTCGCCAGCTAGATGACCTTAATATGGCATGGAAAC 4082
QY 1285 TTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTG 1344
Db 4083 TTTTACAGGTTTCTGTGGATGATCGCCTTAAACAGCTTCAGGAAGCCCAAGATTTTG 4142
QY 1345 GTCCAGCATCTCAGCACTTTCTTTCCACGCTGTCTCCAGGGTCCCTGGGAGAGAGCCATCT 1404
Db 4143 GACCATCCTCTCAGCATTTTCTCTCTACGTCAGTCAGCTGCCGTGGCAAGATCCATTT 4202
QY 1405 CGCCAAACAAAGTCCCTTACTATATCAACCCAGGACTCAAAACAACTTGTCTGGGACCATC 1464
Db 4203 CACATAATAAGTGCCTTATTACATCAACCATCAACATCAACACAGACCCTGTGGGACCATC 4262
QY 1465 CCAAAATGACAGAGCTCTACCAGTCTTTTAGCTGACCTGAATAATGTACGATTTCTCAGCTT 1524
Db 4263 CTAAATGACCGAACTCTTCAATCCCTTGTGACCTGAATAATGTACGTTTCTGCTCCT 4322
QY 1525 ATAGGACTGCCATGAATCCGAACTCCGAAGACTGCAGAAAGCCCTTTGCTTGGATCTCTTGAGCC 1584
Db 4323 ACCGTACAGCAATCAAAATCCGAAGACTACAAAAGCACTATGTTTGGATCTCTTAGAGT 4382
QY 1585 TGTGAGCTGCATGTGATGCCCTTGGACCAGCACAACTCAAGCAAAATGACAGCCCATGG 1644
Db 4383 TGAGTACAACAATGAATTTTCAAACAGCACAAAGTTGAACCAAAATGACAGCTCCTCA 4442
QY 1645 ATATCCTGCAGATTATTAATGTTTGAACCACTATTTATGACCGCTGGAGCAAGAGCACA 1704
Db 4443 GTGTTCCAGATGTCACTCAACTGTCTGACAACTTATGATGAGCTTGAGCAATGCATA 4502
QY 1705 ACAATTTGGTCAACGTCCTCTCTGCGTGGATGATGTCTGAACTGGCTGCTGAATGTTT 1764
Db 4503 AGGACCTGGTCAACGTTCCACTCTGTGTTGATGATGTCTCAATTTGCTCAATGTCT 4562
QY 1765 ATGATACGGGACGACAGGGAGGATCCGTGTCTGCTTTTAAACTGGCATCATTTCCC 1824
Db 4563 ATGACACGGGTGCACTGGAATAATTAGAGTCAGAGTCTGAAGATTGATTATGCTCTC 4622
QY 1825 TGTGTAAGCACATTTTGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAA 1884
Db 4623 TCTCCAAAGGTCTCTTGAAGAAAATAACAGATATCTCTTTAAGGAAGTTGCGGGCCGA 4682
QY 1885 CAGGATTTTGTGACCAAGCGCAGGCTGGGCCCTCTTCTGCATGATTTCTATCCAAATTCCA 1944
Db 4683 CAGAAATGTGTGACCAAGGAGCAGCTGGGCCCTGTTACTTCAATGATGCCATCCAGATCCC 4742
QY 1945 GACAGTTGGTGAGTTGCATCCCTTTGGGGGCGAGTAACATTGAGCCAAGTGTCCGGA 2001
Db 4743 GGCAGCTAGGTGAAGTAGCAGCTTTTGGAGGCGAGTAATATTGAGCCTAGTGTTCGCA 4799

RESULT 6

US-09-091-501B-9
; Sequence 9, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B

; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 10320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(10312)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Full length
; OTHER INFORMATION: utrophin construct
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (724)..(758)
; OTHER INFORMATION: Precise residue is left open
US-09-091-501B-9

Query Match 24.5%; Score 490.6; DB 4; Length 10320;
Best Local Similarity 67.6%; Pred. No. 1.4e-141;
Matches 688; Conservative 0; Mismatches 329; Indels 0; Gaps 0;
QY 985 AAAAGAGTACAGCACAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGC 1044
Db 8058 AAAAGCAAGTGGACAAGGCATTGGAGAACTCAGAGACCTGCAGGGAGCTATGGATGACC 8117
QY 1045 TGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATC 1104
Db 8118 TGGACGCTGACATGAAGGAGGCAGAGTCCGTGCGGAATGGCTGGAAGCCCGTGGGAGACT 8177
QY 1105 TCCTCATTTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTG 1164
Db 8178 TACTCATTTGACTCGCTGCAGGATCACATTGAAAAAATCATGGCATTTAGAGAAGAAATTG 8237
QY 1165 CGCCTCTGAAAGAGAACGTCAGCCACGTCATCAATGACCTTGTCTGCCAGCTTACCACCTTTGG 1224
Db 8238 CACCAATCAACTTTAAAGTTAAACCGTGAATGATTTATCCAGTCAGCTGTCTCCACTTG 8297
QY 1225 GCATTCAGCTCTCACCGTATAAACCTCAGCACTCTGGAAGACCTTGAACACCCAGATGGAAGC 1284
Db 8298 ACCTGCATCCCTCTCTAAAGATGTCTCGCCAGCTAGATGACCTTAATATATCGATGGAAC 8357
QY 1285 TTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTG 1344
Db 8358 TTTTACAGGTTTCTGTGGATGATCGCCTTAAACAGCTTCAGGAAGCCACAGAGATTTTG 8417
QY 1345 GTCCAGCATCTCAGCACTTTCTTTCCACGCTGTCTCCAGGGTCCCTGGGAGAGAGCCATCT 1404
Db 8418 GACCATCCTCTCAGCATTTTCTCTCTACGTCAGTCCAGCTGCCGTGGCAAGATCCATTT 8477
QY 1405 CGCCAAACAAAGTCCCTTACTATATCAACCCAGGACTCAAAACAACTTGTCTGGGACCATC 1464
Db 8478 CACATAATAAGTGCCCTATTACATCAACCATCAAAACACAGACCACCTGTGGGACCATC 8537
QY 1465 CCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTACAGATTCTCAGCTT 1524
Db 8538 CTAAATGACCGAACTCTTTCAATCCCTTGTGCTGACCTGAATAATGTACGTTTCTGCTCCT 8597
QY 1525 ATAGGACTGCCATGAAACTCCGAAGACTGCAGAGAGCCCTTTTGTGGATCTCTTGAGCC 1584
Db 8598 ACCGTACAGCAATCAAAATCCGAAGACTACAAAAGCACTATGTTGGATCTCTTAGAGT 8657
QY 1585 TGTGAGCTGCATGTGATGCCCTTGGACCAGCACAACTCAAGCAAAATGACCCAGCCATGG 1644

Db 8658 TGAGTACAACAAATGAATTTTCAAACAGCACAAAGTTGAACCAAAATGACCAGCTCCTCA 8717
QY 1645 ATATCCTGCAGATTATTAATTGTTTGACCACCTATTTATGACCGCTCGAGCAAGGCACA 1704
Db 8718 GTGTTCCAGATGTCACTCACTGTCTGACAACAACCTATGATGGACTTGACCAATGCATA 8777
QY 1705 ACAATTGGTCAACGTCCTCTCTCGTGGGATATGTCTGAACTGGCTGCTGAATGTTT 1764
Db 8778 AGGACCTGGTCAACGTTCCACTCTGTGTTGATATGTCTCAATTGGTTGCTCAATGTCT 8837
QY 1765 ATGATACGGGACGAACAGGAGGATCCGTGTCCTGCTTTTAAAACTGGCATCTTCCC 1824
Db 8838 ATGACACGGGTGCAACTGGAATAATTAGAGTGCAGAGTCTGAAGATTGGATTAAATGCTC 8897
QY 1825 TGTGTAAGCACATTTTGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAA 1884
Db 8898 TCTCCAAAGGTCTCTTTGGAAGAAAATACAGATATCTCTTTAAGGAAGTTGCGGGCCGA 8957
QY 1885 CAGGATTTGTGACCAAGCGCAGGCTGGGCTCTTCTGCATGATTCATCCAAATCCAA 1944
Db 8958 CAGAAATGTGTGACCAAGCGCAGCTGGGCTGTTACTTCATGATGCCATCCAGATCCGCC 9017
QY 1945 GACAGTTGGTGAAGTTGCATCCTTTGGGGGAGTAACATTGAGCCAAAGTGTCCGA 2001
Db 9018 GGCAGCTAGGTGAAGTAGCAGCTTTTGGAGGCAAGTAATATTGAGCCTAGTGTTCGCA 9074

RESULT 7

US-09-091-501B-5
; Sequence 5, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Uterophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Rattus sp.
US-09-091-501B-5

Query Match 4.0%; Score 79.4; DB 4; Length 200;
Best Local Similarity 64.3%; Pred. No. 9e-15;
Matches 119; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 464 CCTAAAACGCCAAGTACAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGCAAGT 523
Db 16 CTGCAAAACCTGCTTGAAGAACAATAAAAGTTTGCAAGTAGCACCTCGAAGCTGAGCAGGT 75
QY 524 CAGGTCAATTCTCTCACTCACATGGTGGTGGTAGTGATGAATCTAGTGGAGATCACGC 583
Db 76 GAAGTGAATTCTTAACTCATATGGTGGTGAATTTGGATGAAAACAGTGGGAGAGCGC 135
QY 584 AACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAG 643
Db 136 CACAGTGTGTTTGGAGAGATCAGTTACAGAACTGGGTGAGCGCTGGACAGCTGTATGCCG 195
QY 644 ATGGA 648
Db 196 CTGGA 200

RESULT 8

US-09-091-501B-4
; Sequence 4, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Uterophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Mus sp.
US-09-091-501B-4

Query Match 3.9%; Score 78.6; DB 4; Length 200;
Best Local Similarity 62.4%; Pred. No. 1.6e-14;
Matches 123; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY 452 TGATCTTGAAGACCTTAAACGCCCAAGTACAACAACATAAGGTGCTTCAAGAAGATCTAGA 511
Db 4 TGACCTGCCCTCCCTGCAGAAGCTGCTTCAAGAACAATAAAAGTTTGCAAAATGACCTTGA 63
QY 512 ACAAGAACAAGTCAGGTCAATTCTCTCACTCACATGGTGGTGGTAGTTGATGAATCTAG 571
Db 64 ACCTGAACAGGTGAAGTAAATTCCTTAACCTCACATGGTGGTGAATTTGGATGAAACAG 123
QY 572 TGGAGATCACGCACTGCTGCTTTGGAAGAACAACCTTAAGGTATTGGGAGATCGATGGGC 631
Db 124 TGGGAGAGTGCACAGCTCTTCTGGAAGATCAGTTACAGAAACTGGGTGAGCGCTGGAC 183
QY 632 AAACATCTGTAGATGGA 648
Db 184 AGCTGTATGCCGCTGGA 200

RESULT 9

US-09-091-501B-6
; Sequence 6, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Uterophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 200

```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-091-501B-6

Query Match      3.9%; Score 78.6; DB 4; Length 200;
Best Local Similarity 62.4%; Pred. No. 1.6e-14;
Matches 123; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 452 TCATCTTGAAGACCTAAACGCCCAAGTACAACAACATAAGGTGCTTCAAGAAGATCTAGA 511
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4 TCATGTGAAATCTCTACAAAAGCTGCTAGAGAACAATAAAAGTTTGCAAAGTGATCTTGA 63

QY 512 ACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGTTGATGAATCTAG 571
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 64 GGCTGAACAGGTGAAAGTAAATTCACAACTCACATGGTGGTTCATTGTTGATGAAAACAG 123

QY 572 TGGAGATCACGCAACTGCTGCTTTTGGGAAGAACAACTTAAGGTATTGGGAGATCGATGGGC 631
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124 TGGTGAGAGCGCTACAGCTATCCTAGAGACCAGTACAGAAACTTGGTGAGCGCTGGAC 183

QY 632 AAACATCTGTAGATGA 648
Db 184 AGCAGTATGCCGTTGA 200

RESULT 10
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14
```

```

Query Match      3.7%; Score 74.6; DB 1; Length 7218;
Best Local Similarity 5.6%; Pred. No. 3.4e-12;
Matches 23; Conservative 236; Mismatches 150; Indels 0; Gaps 0;

QY 113 TTCTAATGATGTGGAACTGGTGAAAGACCAGTTTCATCTATCATGAGGGGTACATGATGG 172
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1474 TATCTATGCAAGTAGTTAAAGAGATAGAAGAAATTGGTACRRRRRRRRRRRRRRRR 1415

QY 173 ATTTGACAGCCCATCAGGCGGGTGGTAATATTTACAATTTGGAAGTAGTGATTG 232
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1414 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1355

QY 233 GAACAGGAAAATTATCAGAAGATGAAGAACTGAAGTACAAGACGAGATGAATCTCTAA 292
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1354 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1295

QY 293 ATTCAAGATGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTACATAGAG 352
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1294 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1235

QY 353 TTTTAATGGATCTCCAGAATCGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGA 412
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1234 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1175

QY 413 AGAAAGAACAAAGGAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTGAAGACCTAAACG 472
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1174 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1115

QY 473 CCAAGTACAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAACAA 521
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1114 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRA 1066

RESULT 11
US-09-668-313A-10
; Sequence 10, Application US/09668313A
; Patent No. 6503756
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Susan M. Freier
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF SYNTAXIN 4 INTERACTING PROTEIN EXPRESSION
; FILE REFERENCE: RTS-0127
; CURRENT APPLICATION NUMBER: US/09/668,313A
; CURRENT FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 247
; SEQ ID NO 10
; LENGTH: 2574
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (218)...(1891)
US-09-668-313A-10

Query Match      2.2%; Score 44.2; DB 4; Length 2574;
Best Local Similarity 54.7%; Pred. No. 0.0048;
Matches 88; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 1389 TGGGAGAGAGCCATCTCGCCAAACAAGTGCCCTACTATATCAACCACGAGACTCAAACA 1448
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1733 TGGGAGGAAGCTTACACAGCAGATGGGAATCAAGTACTTCATCAACCACGTCACACAGACC 1792

QY 1449 ACTTGCTGGGACCATCCCAAAATGACAGAGTCTTACCAGTCTTTAGTGCCTGAATAAT 1508
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1793 ACGTCTGGATCCACCCCGTGATGAGCGGCCCTGAACCTGTCTCTGTGAGAGGAGTGAA 1852

QY 1509 GTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAG 1549
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1853 GAGGACTGTCCAGAGAGCTAACAGACCCCGAAAAGCTGATG 1893
```


RESULT 12

```
US-09-107-532A-1186
; Sequence 1186, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 1186:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1179 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...1179
; SEQUENCE DESCRIPTION: SEQ ID NO: 1186:
US-09-107-532A-1186
Query Match 2.1%; Score 42.8; DB 4; Length 1179;
Best Local Similarity 49.5%; Pred. No. 0.0076;
Matches 110; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
Qy 353 TTTTAATGGATCTCCAGAAATCGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGA 412
Db 585 TGTGATCCAGCAGTTGAGCAAGAATAAAGATTGATCGATCAGACAAAGAAAAATGG 644
Qy 413 AGAAGAACAAAGGAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTGAAGACCTAAAACG 472
Db 645 AGATACGATCGGAGGAATTGAGAAAGTGCTCGTTGGAGGCGTTCAGCTGGATTAGGAAG 704
Qy 473 CCAAGTACAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGACAGTCAGGGTCAA 532
Db 705 CTACGTACAATGGGACACGAAGCTAGATGCCAAAATCGCACAAAGCTGTGTTAGTATCAA 764
Qy 533 TTCTCTCACTCACATGGTGGTGGTAGTTGATGAATCTAGTGG 574
Db 765 TGCCTTTAAAGCGGTAGAATTGGGGTCGGATTCACTTCTGG 806
```

RESULT 13

```
US-09-620-312D-69
; Sequence 69, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 69
; LENGTH: 1690
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (128)..(1522)
US-09-620-312D-69
```

```
Query Match 2.1%; Score 42.8; DB 4; Length 1690;
Best Local Similarity 47.9%; Pred. No. 0.0097;
Matches 156; Conservative 0; Mismatches 167; Indels 3; Gaps 1;
Qy 1016 CCAGGAACCTCAAGAGCCACGGATGAGCTGAGCCTCAAGCTGCGCCAAAGCTGAGGTGAT 1075
Db 55 CAAGGAGTTGCACCCAGGTGGCGCACGACCTGGACGACGAGCTGGCATGGGTTTCAGGAGCG 114
Qy 1076 CAAGGGA---TCCTGGCAGCCCGTGGCGGATCTCTCATTGACTCTCTCCAAGATCACCT 1132
Db 115 GCTGCCACTGGCCATGCAGACAGAGCGAGGCAACGGTTTGCAGGGCGGTCCAGCAGCAT 174
Qy 1133 CGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGT 1192
Db 175 CAAAAAGAACCAAGGCGCTGCGGCGGAGATCCAGGCGCATGGGCGCGCTGGAGGAGGT 234
Qy 1193 CAATGACCTTGCTCGCAGCTTACCACCTTTGGGCATTTCAGCTCTCACCGTATAACCTCAG 1252
Db 235 GCTGGAGCGCGGCGGCGCTGGCGTCTGCGCAGCCCGGAGGAGGAGGAGGTGCGCGC 294
Qy 1253 CACTCTGGAAGACCTGAACACCAAGATGGAAGCTTTCAGAGTGGCCCGTCGAGGACCGAGT 1312
Db 295 GGGCCTGGAGCAGCTGCAGAGCGCCTGGGCGGAGCTGCGGAGGCTGCCGAGCGACGGCA 354
Qy 1313 CAGCAGCTGCATGAAGCCACAGGG 1338
Db 355 GCAGGTGCTGGACGCCGCTTCCAGG 380
```

RESULT 14

US-09-368-590-1
; Sequence 1, Application US/09368590
; Patent No. 6187563
; GENERAL INFORMATION:
; APPLICANT: Solimena, Michele
; TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
; FILE REFERENCE: 101918-200 (OCR-941)
; CURRENT APPLICATION NUMBER: US/09/368,590
; CURRENT FILING DATE: 1999-08-04
; EARLIER FILING DATE: 1998-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 7812
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(6879)
; NAME/KEY: unsure
; LOCATION: (100)...(102)
; NAME/KEY: unsure
; LOCATION: (1021)...(1023)
; NAME/KEY: unsure
; LOCATION: (2266)...(2268)
US-09-368-590-1

Query Match 2.1%; Score 42.8; DB 3; Length 7812;
Best Local Similarity 47.9%; Pred. No. 0.028;
Matches 156; Conservative 0; Mismatches 167; Indels 3; Gaps 1;
QY 1016 CCAGGAAGTTCAGAGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGAT 1075
Db 3639 CAAGGAGTTGCACCGTGGCGGCACGACCTGGACGACGAGCTGGCATGGGTTTCAGGAGCG 3698
QY 1076 CAAGGGA---TCCTGGCAGCCCGTGGCGATCTCTCTATTGACTCTCTCCAAGATCACCT 1132
Db 3699 GCTGCCACTGGCCATGCAGACAGAGCGAGGCAACGGTTTGCAGGCGGTCCAGCAGCAT 3758
QY 1133 CGAGAAAGTCAAGGCACTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCAGT 1192
Db 3759 CAAAAGAACCCAGGCGCTGCGGCGGAGATCCAGGCGCATGGGCCGCGCCTGGAGGAGGT 3818
QY 1193 CAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTTCAGCTCTACCGGTATAACCTCAG 1252
Db 3819 GCTGGAGCGCGCGGCGCGCTGGCGTCTGCTGCCAGCCCGAGGAGGAGGAGGAGGAGG 3878
QY 1253 CACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAAGTGGCGCTGAGGACCGAGT 1312
Db 3879 GGGCCTGGAGCAGCTGCAGAGCGCTGGGCCGAGCTGGGGAGGCTGCCGAGCGGCA 3938
QY 1313 CAGGCAGCTGCATGAAGCCACAGGG 1338
Db 3939 GCAGGTGCTGGACGCCGCCCTTCCAGG 3964

RESULT 15
US-08-857-076-44/c
; Sequence 44, Application US/08857076C
; Patent No. 6225120
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Kimura, Koutarou
; APPLICANT: Patterson, Garth
; APPLICANT: Ogg, Scott
; APPLICANT: Paradis, Suzanne
; APPLICANT: Tissenbaum, Heidi
; APPLICANT: Morris, Jason
; APPLICANT: Kowee, Allison
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; IMPAIRED GLUCOSE TOLERANCE CONDITIONS

FILE REFERENCE: 00786/351001
; CURRENT APPLICATION NUMBER: US/08/857,076C
; CURRENT FILING DATE: 1997-05-15
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 2704
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-08-857-076-44

Query Match 2.1%; Score 42.4; DB 3; Length 2704;
Best Local Similarity 48.4%; Pred. No. 0.018;
Matches 118; Conservative 0; Mismatches 126; Indels 0; Gaps 0;
QY 196 GTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGAT 255
Db 2116 CCTAGACATTTGCTACCGGGGAAAAAAGAGGAAAAATTAAAGAAATTACAAGAGAAAT 2057
QY 256 GAAGAAACTGAAGTACAAGAGCAGATGAATCTCTTAATTTCAAGATGGGAATGCCTCAGG 315
Db 2056 CAAGAAATCAAAATTTTGAATAATTCAATTAATAAAGACAGATTGGGGAATGTGTTT 1997
QY 316 GTAGCTAGCATGGAATAAACAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCGA 375
Db 1996 GCAGTTTGTGATTGCGAAACAAATTTACAGCGAAAAAGATGGGATAGAGGTAGCATTTAA 1937
QY 376 AACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAAACAAAGGAAAAATGGAGG 435
Db 1936 GAGTGAAGAACCTGGATGACGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGT 1877
QY 436 AAGA 439
Db 1876 AATA 1873

Search completed: February 2, 2004, 06:16:03
Job time : 125.697 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 23:51:42 ; Search time 655.123 Seconds
(without alignments)
11131.886 Million cell updates/sec

Title: US-09-845-416-14_COPY_1000_3000
Perfect score: 2001
Sequence: 1 ggcagttcattgatggagag.....cattgagccaagtgtccgga 2001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2434939 seqs, 1622278265 residues
Total number of hits satisfying chosen parameters: 4869878

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

| | |
|-----|--|
| 1: | /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:* |
| 2: | /cgn2_6/ptodata/2/pubpna/pct_NEW_PUB.seq:* |
| 3: | /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:* |
| 4: | /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:* |
| 5: | /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:* |
| 6: | /cgn2_6/ptodata/2/pubpna/pctUS_PUBCOMB.seq:* |
| 7: | /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:* |
| 8: | /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:* |
| 9: | /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:* |
| 10: | /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:* |
| 11: | /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:* |
| 12: | /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:* |
| 13: | /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:* |
| 14: | /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:* |
| 15: | /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:* |
| 16: | /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:* |
| 17: | /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:* |
| 18: | /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:* |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID | Description |
|------------|--------|-------|--------|----|------------------|-------------------|
| 1 | 2001 | 100.0 | 3446 | 13 | US-09-845-416-14 | Sequence 14, Appl |
| 2 | 1990 | 99.5 | 4414 | 13 | US-09-845-416-32 | Sequence 32, Appl |
| 3 | 1820.8 | 91.0 | 5417 | 13 | US-10-149-736-39 | Sequence 39, Appl |
| 4 | 1593 | 79.6 | 3510 | 13 | US-09-845-416-12 | Sequence 12, Appl |
| 5 | 1593 | 79.6 | 4476 | 13 | US-09-845-416-31 | Sequence 31, Appl |
| 6 | 1263 | 63.1 | 3858 | 13 | US-09-845-416-9 | Sequence 9, Appl |
| 7 | 1263 | 63.1 | 4825 | 13 | US-09-845-416-29 | Sequence 29, Appl |
| 8 | 1263 | 63.1 | 4848 | 13 | US-09-845-416-35 | Sequence 35, Appl |
| 9 | 1263 | 63.1 | 5060 | 13 | US-09-845-416-36 | Sequence 36, Appl |
| 10 | 1254 | 62.7 | 5339 | 13 | US-10-149-736-40 | Sequence 40, Appl |
| 11 | 1251.2 | 62.5 | 3531 | 13 | US-09-845-416-10 | Sequence 10, Appl |
| 12 | 1251.2 | 62.5 | 4498 | 13 | US-09-845-416-30 | Sequence 30, Appl |
| 13 | 1245 | 62.2 | 4182 | 13 | US-09-845-416-2 | Sequence 2, Appl |
| 14 | 1245 | 62.2 | 5149 | 13 | US-09-845-416-27 | Sequence 27, Appl |
| 15 | 1132.2 | 56.6 | 5462 | 13 | US-10-149-736-41 | Sequence 41, Appl |

| | | | | | | |
|----|--------|------|-------|----|---------------------|--------------------|
| 16 | 1112 | 55.6 | 3999 | 13 | US-09-845-416-6 | Sequence 6, Appl |
| 17 | 1112 | 55.6 | 4966 | 13 | US-09-845-416-28 | Sequence 28, Appl |
| 18 | 1112 | 55.6 | 4990 | 13 | US-09-845-416-34 | Sequence 34, Appl |
| 19 | 1002.6 | 50.1 | 1821 | 13 | US-09-845-416-13 | Sequence 13, Appl |
| 20 | 1002.6 | 50.1 | 2169 | 13 | US-09-845-416-4 | Sequence 4, Appl |
| 21 | 1002.6 | 50.1 | 8689 | 13 | US-10-149-736-42 | Sequence 42, Appl |
| 22 | 1002.6 | 50.1 | 11058 | 13 | US-09-845-416-1 | Sequence 1, Appl |
| 23 | 1002.6 | 50.1 | 11443 | 13 | US-10-149-736-44 | Sequence 44, Appl |
| 24 | 1002.6 | 50.1 | 12057 | 13 | US-10-149-736-47 | Sequence 47, Appl |
| 25 | 1002.6 | 50.1 | 13957 | 10 | US-09-782-378A-22 | Sequence 22, Appl |
| 26 | 1002.6 | 50.1 | 13957 | 10 | US-09-880-107-2384 | Sequence 2284, Ap |
| 27 | 1002.6 | 50.1 | 13957 | 13 | US-10-149-736-1 | Sequence 1, Appl |
| 28 | 1002.6 | 50.1 | 14082 | 13 | US-10-341-434-108 | Sequence 108, Appl |
| 29 | 1001 | 50.0 | 1434 | 13 | US-09-845-416-15 | Sequence 15, Appl |
| 30 | 985.8 | 49.3 | 1991 | 13 | US-09-845-416-3 | Sequence 3, Appl |
| 31 | 869.8 | 43.5 | 13815 | 13 | US-10-149-736-2 | Sequence 2, Appl |
| 32 | 665 | 33.2 | 1667 | 13 | US-09-845-416-7 | Sequence 7, Appl |
| 33 | 506.6 | 25.3 | 11096 | 13 | US-10-149-736-4 | Sequence 4, Appl |
| 34 | 490.6 | 24.5 | 10302 | 10 | US-09-782-378A-23 | Sequence 23, Appl |
| 35 | 490.6 | 24.5 | 10302 | 13 | US-10-149-736-3 | Sequence 3, Appl |
| 36 | 481.8 | 24.1 | 16531 | 13 | US-10-101-510-667 | Sequence 667, App |
| 37 | 457 | 22.8 | 887 | 13 | US-10-149-736-35 | Sequence 35, Appl |
| 38 | 350 | 17.5 | 1340 | 13 | US-09-845-416-11 | Sequence 11, Appl |
| 39 | 327 | 16.3 | 327 | 13 | US-10-149-736-8 | Sequence 8, Appl |
| 40 | 324 | 16.2 | 324 | 13 | US-10-149-736-33 | Sequence 33, Appl |
| 41 | 322 | 16.1 | 333 | 13 | US-10-149-736-10 | Sequence 10, Appl |
| 42 | 321 | 16.0 | 333 | 13 | US-10-149-736-9 | Sequence 9, Appl |
| 43 | 216 | 10.8 | 216 | 13 | US-10-149-736-34 | Sequence 34, Appl |
| 44 | 180 | 9.0 | 476 | 9 | US-09-864-761-15766 | Sequence 15766, A |
| 45 | 160.2 | 8.0 | 256 | 9 | US-09-864-761-21956 | Sequence 21956, A |

ALIGNMENTS

RESULT 1
US-09-845-416-14
; Sequence 14, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 3446
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-14

| | | | | | | | |
|-----------------------|---------|---|-------|------------|-----|--------|-------|
| Query Match | 100.0%; | Score | 2001; | DB | 13; | Length | 3446; |
| Best Local Similarity | 100.0%; | Pred. No. | 0; | | | | |
| Matches | 2001; | Conservative | 0; | Mismatches | 0; | Indels | 0; |
| | | | | | | | |
| Qy | 1 | GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAA | 60 | | | | |
| | | | | | | | |
| Db | 1000 | GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAA | 1059 | | | | |
| | | | | | | | |
| Qy | 61 | GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGCAAGCACAGGAGAGATTCTTAAT | 120 | | | | |
| | | | | | | | |
| Db | 1060 | GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGCAAGCACAGGAGAGATTCTTAAT | 1119 | | | | |
| | | | | | | | |
| Qy | 121 | GATGTGGAAGTGGTGAAGACACAGTTTTCATCTCATGAGGGGTACATGATGGATTGACA | 180 | | | | |
| | | | | | | | |
| Db | 1120 | GATGTGGAAGTGGTGAAGACACAGTTTTCATCTCATGAGGGGTACATGATGGATTGACA | 1179 | | | | |
| | | | | | | | |
| Qy | 181 | GCCCATCAGGGCCGGGTGGTAAATATCTACAATTGGGAAGTAAGCTGATTGGAACAGGA | 240 | | | | |

Db 1180 GCCCATCAGGCGCGGTGGTAAATATCTACAATTGGGAAGTAAGCTGATGGAAACAGGA 1239
QY 241 AAATTATCAGAAGATGAAGAAAACCTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 300
Db 1240 AAATTATCAGAAGATGAAGAAAACCTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 1299
QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGA AAAACAAAAGCAATTTACATAGAGTTTAAATG 360
Db 1300 TGGGAATGCCTCAGGGTAGCTAGCATGGA AAAACAAAAGCAATTTACATAGAGTTTAAATG 1359
QY 361 GATCTCCAGAAATCGAAACTGAAAGAGTTGAATGACTGGCTTAACAAAACAGAAAGAA 420
Db 1360 GATCTCCAGAAATCGAAACTGAAAGAGTTGAATGACTGGCTTAACAAAACAGAAAGAA 1419
QY 421 CAAGGAAAATGGAGGAAGAGCCTCTTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTAC 480
Db 1420 CAAGGAAAATGGAGGAAGAGCCTCTTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTAC 1479
QY 481 AACAAACATAAGGTGCTTCAAGAAAGATCTAGAACAAAGAAACAAGTCAGGTCATCTCTCA 540
Db 1480 AACAAACATAAGGTGCTTCAAGAAAGATCTAGAACAAAGAAACAAGTCAGGTCATCTCTCA 1539
QY 541 CTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAG 600
Db 1540 CTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAG 1599
QY 601 AACAACTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCT 660
Db 1600 AACAACTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCT 1659
QY 661 GGGTCTCTTTTACAAGACATCCCTTCTCAAAATGGCAACGTCTTACTGAAGAACAGTGCCTTT 720
Db 1660 GGGTCTCTTTTACAAGACATCCCTTCTCAAAATGGCAACGTCTTACTGAAGAACAGTGCCTTT 1719
QY 721 TTAGTGCATGGCTTTCAGAAAAAGAAAGATGCAAGTGAACAAGATTCACAAACTGGCTTTA 780
Db 1720 TTAGTGCATGGCTTTCAGAAAAAGAAAGATGCAAGTGAACAAGATTCACAAACTGGCTTTA 1779
QY 781 AAGATCAAAATGAATGTTATCAAGTCTTCAAAAATGGCCGCTTTTAAAGCGGATCTAG 840
Db 1780 AAGATCAAAATGAATGTTATCAAGTCTTCAAAAATGGCCGCTTTTAAAGCGGATCTAG 1839
QY 841 AAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACAC 900
Db 1840 AAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACAC 1899
QY 901 TGAAGAATAAGTCAGTGACCCAGAAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTT 960
Db 1900 TGAAGAATAAGTCAGTGACCCAGAAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTT 1959
QY 961 GGGATAAATTTAGTCCAAAAACTTGAAAAAGAGTACAGCACAGACCCCTTGAAAGACTCCAGG 1020
Db 1960 GGGATAAATTTAGTCCAAAAACTTGAAAAAGAGTACAGCACAGACCCCTTGAAAGACTCCAGG 2019
QY 1021 AACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGG 1080
Db 2020 AACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGG 2079
QY 1081 GATCCTGGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAG 1140
Db 2080 GATCCTGGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAG 2139
QY 1141 TCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTGAGCCACGTCATGACC 1200
Db 2140 TCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTGAGCCACGTCATGACC 2199
QY 1201 TTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTGG 1260
Db 2200 TTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTGG 2259
QY 1261 AAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGCTCGAGGACCGAGTCAGGCAGC 1320

Db 2260 AAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGC 2319
QY 1321 TGCATGAAGCCCAACAGGGACTTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCC 1380
Db 2320 TGCATGAAGCCCAACAGGGACTTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCC 2379
QY 1381 AGGTCCTCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACACCAGAGA 1440
Db 2380 AGGTCCTCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACACCAGAGA 2439
QY 1441 CTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACC 1500
Db 2440 CTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACC 2499
QY 1501 TGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGG 1560
Db 2500 TGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGG 2559
QY 1561 CCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGTCATGTGATGCCCTTGGACCAGCAACACC 1620
Db 2560 CCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGTCATGTGATGCCCTTGGACCAGCAACACC 2619
QY 1621 TCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTT 1680
Db 2620 TCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTT 2679
QY 1681 ATGACCGCTGGAGCAAGAGACAAACAATTTGGTCAACGTCCCTCTCTGCGTGGATATGT 1740
Db 2680 ATGACCGCTGGAGCAAGAGACAAACAATTTGGTCAACGTCCCTCTCTGCGTGGATATGT 2739
QY 1741 GTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCCGTGTCTCTGT 1800
Db 2740 GTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCCGTGTCTCTGT 2799
QY 1801 CTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGATACC 1860
Db 2800 CTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGATACC 2859
QY 1861 TTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGCGCTGGGCCCTCCTTC 1920
Db 2860 TTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGCGCTGGGCCCTCCTTC 2919
QY 1921 TGCAATGATTTCTATCCAAATTCCAAGACAGATTGGGTGAAAGTTGCATCCTTTGGGGGCAGTA 1980
Db 2920 TGCAATGATTTCTATCCAAATTCCAAGACAGATTGGGTGAAAGTTGCATCCTTTGGGGGCAGTA 2979
QY 1981 ACATTGAGCCAAAGTGTCCGGA 2001
Db 2980 ACATTGAGCCAAAGTGTCCGGA 3000

RESULT 2

US-09-845-416-32
; Sequence 32, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DB1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 4414
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-32

Query Match 99.5%; Score 1990; DB 13; Length 4414;

| Best Local Similarity 100.0%; Pred. No. 0; | | | | | | | | | | | | | | |
|---|------|---------|---------|----------|---------|----------|---------|---------|----------|-----------------|-----------|------------|-------------|--------------|
| Matches 2001; Conservative 0; Mismatches 0; Indels 1; Gaps 1; | | | | | | | | | | | | | | |
| QY | 1 | GGCAGTT | CATTGAT | GAGAGT | GAAGTAA | ACCTGG | ACCGTT | TATCAAA | CAGCTTT | TAGAAGAA 60 | | | | |
| Db | 1757 | GGCAGTT | CATTGAT | GAGAGT | GAAGTAA | ACCTGG | ACCGTT | TATCAAA | CAGCTTT | TAGAAGAA 1816 | | | | |
| QY | 61 | GTATTAT | CGTGGCT | TCTTTCT | GTGAGG | ACACATTG | CAAGC | ACAAGG | AGAGATTT | TCTTAAT 120 | | | | |
| Db | 1817 | GTATTAT | CGTGGCT | TCTTTCT | GTGAGG | ACACATTG | CAAGC | ACAAGG | AGAGATTT | TCTTAAT 1876 | | | | |
| QY | 121 | GATGTGA | AGTGGT | GAAAGACC | AGTTTCA | TACTCAT | GAGGGT | TACATG | ATGGATT | TGACA 180 | | | | |
| Db | 1877 | GATGTGA | AGTGGT | GAAAGACC | AGTTTCA | TACTCAT | GAGGGT | TACATG | ATGGATT | TGACA 1936 | | | | |
| QY | 181 | GCCCAT | CAGGCG | GGTTGGT | TAATATT | CTACA | ATTGGGA | AGTAAG | CTGATT | TGGAACAGGA 240 | | | | |
| Db | 1937 | GCCCAT | CAGGCG | GGTTGGT | TAATATT | CTACA | ATTGGGA | AGTAAG | CTGATT | TGGAACAGGA 1996 | | | | |
| QY | 241 | AAATTAT | CAGAAG | ATGAAGAA | ACTGAAG | TACAAG | AGCAGAT | GAATCT | CCTAA | ATTCAAGA 300 | | | | |
| Db | 1997 | AAATTAT | CAGAAG | ATGAAGAA | ACTGAAG | TACAAG | AGCAGAT | GAATCT | CCTAA | ATTCAAGA 2056 | | | | |
| QY | 301 | TGGGAAT | GCCTC | AGGTPAG | CTAGCAT | GGA | AAAAA | CAAGCA | ATTTAC | ATAGAGTTTAA | TG 360 | | | |
| Db | 2057 | TGGGAAT | GCCTC | AGGTPAG | CTAGCAT | GGA | AAAAA | CAAGCA | ATTTAC | ATAGAGTTTAA | TG 2116 | | | |
| QY | 361 | GATCTCC | AGAATC | -GAAACT | GAAAGAG | TTGAAT | GACTGG | CTAA | CAAAAA | CAGAAAGAA | 419 | | | |
| Db | 2117 | GATCTCC | AGAATC | AGAACT | GAAAGAG | TTGAAT | GACTGG | CTAA | CAAAAA | CAGAAAGAA | 2176 | | | |
| QY | 420 | ACAAGG | AAATG | GGAAGAG | CCCTCT | TGGAC | CTGATCT | TGAAG | ACCTTAA | ACGCCA | AGTA 479 | | | |
| Db | 2177 | ACAAGG | AAATG | GGAAGAG | CCCTCT | TGGAC | CTGATCT | TGAAG | ACCTTAA | ACGCCA | AGTA 2236 | | | |
| QY | 480 | CAACAAC | ATAAGT | GCTTCA | AGAAGAT | CTAGA | ACAAGAA | CAAGT | CAGGT | CAATTCT | CTC 539 | | | |
| Db | 2237 | CAACAAC | ATAAGT | GCTTCA | AGAAGAT | CTAGA | ACAAGAA | CAAGT | CAGGT | CAATTCT | CTC 2296 | | | |
| QY | 540 | ACTCAC | ATGGT | GGTAG | TTGATGA | ATCTAG | TGGAG | ATCACG | CAACTG | CTGCTTT | GGAA 599 | | | |
| Db | 2297 | ACTCAC | ATGGT | GGTAG | TTGATGA | ATCTAG | TGGAG | ATCACG | CAACTG | CTGCTTT | GGAA 2356 | | | |
| QY | 600 | GAAACA | ACTTAAG | STATTGG | GAGATCG | ATGGG | CAAA | ACATCT | GTAGAT | GACAGAA | CGC 659 | | | |
| Db | 2357 | GAAACA | ACTTAAG | STATTGG | GAGATCG | ATGGG | CAAA | ACATCT | GTAGAT | GACAGAA | CGC 2416 | | | |
| QY | 660 | TGGGTT | CTTTT | TACAAG | ACATCCT | TCTCAA | ATGG | CAACGTC | TCTT | ACTGAAG | ACAGTGC | TT 719 | | |
| Db | 2417 | TGGGTT | CTTTT | TACAAG | ACATCCT | TCTCAA | ATGG | CAACGTC | TCTT | ACTGAAG | ACAGTGC | TT 2476 | | |
| QY | 720 | TTTAGT | GCATGG | CTTT | CAGAAAA | AGAGAT | GAGTGA | ACCAAG | ATTCA | CAACA | ACTGG | CTTT 779 | | |
| Db | 2477 | TTTAGT | GCATGG | CTTT | CAGAAAA | AGAGAT | GAGTGA | ACCAAG | ATTCA | CAACA | ACTGG | CTTT 2536 | | |
| QY | 780 | AAAGAT | CAAAAT | GAAATG | TTATCA | AGTCTT | CAAAAA | CTGG | CCGTTTT | TAAA | AGCGG | ATCTA 839 | | |
| Db | 2537 | AAAGAT | CAAAAT | GAAATG | TTATCA | AGTCTT | CAAAAA | CTGG | CCGTTTT | TAAA | AGCGG | ATCTA 2596 | | |
| QY | 840 | GAAAA | GAAAA | AAAGCA | ATCCAT | GGGCA | AACTG | TATTCA | CTCA | AAACA | AGATCT | TCTTT | CAACA 899 | |
| Db | 2597 | GAAAA | GAAAA | AAAGCA | ATCCAT | GGGCA | AACTG | TATTCA | CTCA | AAACA | AGATCT | TCTTT | CAACA 2656 | |
| QY | 900 | CTGAAG | AATAAG | TCAGT | GACCC | CAGAG | ACGG | AAGCAT | GGCT | GGATA | AACTTT | TGCC | CGGTGT 959 | |
| Db | 2657 | CTGAAG | AATAAG | TCAGT | GACCC | CAGAG | ACGG | AAGCAT | GGCT | GGATA | AACTTT | TGCC | CGGTGT 2716 | |
| QY | 960 | TGGGAT | AATTTAG | TCCAAAA | AACTTG | AAAA | AGAT | TACAC | AGAC | CCCTT | GAA | AGACT | CCAG 1019 | |
| Db | 2717 | TGGGAT | AATTTAG | TCCAAAA | AACTTG | AAAA | AGAT | TACAC | AGAC | CCCTT | GAA | AGACT | CCAG 2776 | |
| QY | 1020 | GAACTT | CAAG | AGCC | ACCG | ATGAG | CTGG | ACCT | CAAG | CTCG | CC | CAAGCT | GAGGT | GATCAAG 1079 |

| | | | | | | | | | | | |
|----|------|--------|--------|-------|--------|---------|--------|--------|--------|---------|-----------------------------------|
| Dd | 2777 | GAACTT | CAAG | AGCC | ACCG | ATGAG | CTGG | ACCT | CAAG | CTGCG | CAAGCTGAGGTGATCAAG 2836 |
| Qy | 1080 | GGATCT | TGGC | AGCC | CGTGG | CGGAT | CTCT | CTAT | TGACT | CTCT | TCCAAGATCACCTCGAGAAA 1139 |
| Dd | 2837 | GGATCT | TGGC | AGCC | CGTGG | CGGAT | CTCT | CTAT | TGACT | CTCT | TCCAAGATCACCTCGAGAAA 2896 |
| Qy | 1140 | GTCAAG | GCAC | TTTC | GAGG | AGAAAT | TGGC | CTCT | GAAAG | AGAA | CGTGAGCCACGTCATGAC 1199 |
| Dd | 2897 | GTCAAG | GCAC | TTTC | GAGG | AGAAAT | TGGC | CTCT | GAAAG | AGAA | CGTGAGCCACGTCATGAC 2956 |
| Qy | 1200 | CTTGCT | CGCC | AGCTT | TACCA | CTTTGG | GCATT | CAGCT | CTCA | CCGTATA | ACCTCAGCACTCTG 1259 |
| Dd | 2957 | CTTGCT | CGCC | AGCTT | TACCA | CTTTGG | GCATT | CAGCT | CTCA | CCGTATA | ACCTCAGCACTCTG 3016 |
| Qy | 1260 | GAAAG | ACCTGA | ACAC | CAGAT | GGAAG | CTTCTG | CAGGT | GGCCGT | CGAGG | ACCGAGTCAGGCAG 1319 |
| Dd | 3017 | GAAAG | ACCTGA | ACAC | CAGAT | GGAAG | CTTCTG | CAGGT | GGCCGT | CGAGG | ACCGAGTCAGGCAG 3076 |
| Qy | 1320 | CTGCAT | GAA | AGCC | CCAC | AGGAC | TTTGGT | CCAG | CATCTC | AGCACT | TTTCTCCACGTCTGTC 1379 |
| Dd | 3077 | CTGCAT | GAA | AGCC | CCAC | AGGAC | TTTGGT | CCAG | CATCTC | AGCACT | TTTCTCCACGTCTGTC 3136 |
| Qy | 1380 | CAGGT | CCCTGG | GAG | AGCC | CATCTC | GC | CAAA | CAAA | AGTGCC | CTACTATATCAACCAACGAG 1439 |
| Dd | 3137 | CAGGT | CCCTGG | GAG | AGCC | CATCTC | GC | CAAA | CAAA | AGTGCC | CTACTATATCAACCAACGAG 3196 |
| Qy | 1440 | ACTCAA | AA | CAAA | CTTGT | GGA | ACC | ATCC | CAAA | ATGAC | AGAGCTCTACCCAGTCTTTAGCTGAC 1499 |
| Dd | 3197 | ACTCAA | AA | CAAA | CTTGT | GGA | ACC | ATCC | CAAA | ATGAC | AGAGCTCTACCCAGTCTTTAGCTGAC 3256 |
| Qy | 1500 | CTGAAT | ATAT | GTCA | AGATT | CTCAG | CTTAT | AGG | ACTG | CAATG | AACTCCGAGACTGCAGAAG 1559 |
| Dd | 3257 | CTGAAT | ATAT | GTCA | AGATT | CTCAG | CTTAT | AGG | ACTG | CAATG | AACTCCGAGACTGCAGAAG 3316 |
| Qy | 1560 | GCCCTT | TGCTT | GGAT | CTCTT | GAG | CCCTG | TCAG | TGAT | GCCTT | TGACCCAGCACAAAC 1619 |
| Dd | 3317 | GCCCTT | TGCTT | GGAT | CTCTT | GAG | CCCTG | TCAG | TGAT | GCCTT | TGACCCAGCACAAAC 3376 |
| Qy | 1620 | CTCAAG | CAAA | ATGAC | AGCC | CAATGG | ATAT | CTCG | AGAT | TATTA | ATTTGTTGACCACTATT 1679 |
| Dd | 3377 | CTCAAG | CAAA | ATGAC | AGCC | CAATGG | ATAT | CTCG | AGAT | TATTA | ATTTGTTGACCACTATT 3436 |
| Qy | 1680 | TATGAC | CGCTT | GAG | CAAG | AGAC | CAACA | ATTTGG | TCAAC | GTCCCT | CTCTGCGTGGATATG 1739 |
| Dd | 3437 | TATGAC | CGCTT | GAG | CAAG | AGAC | CAACA | ATTTGG | TCAAC | GTCCCT | CTCTGCGTGGATATG 3496 |
| Qy | 1740 | TGCTGA | AACTGG | CTGCT | GAAT | GTGTTAT | GAT | AGAG | AGAG | AGAG | ATCCGTGTCTCTG 1799 |
| Dd | 3497 | TGCTGA | AACTGG | CTGCT | GAAT | GTGTTAT | GAT | AGAG | AGAG | AGAG | ATCCGTGTCTCTG 3556 |
| Qy | 1800 | TCTTTT | AAAA | CTGG | CAAT | TTTCC | CTGT | GTAAA | AGCAC | ATTTG | GACCAATTTGGAAGACAAGTACAGATAC 1859 |
| Dd | 3557 | TCTTTT | AAAA | CTGG | CAAT | TTTCC | CTGT | GTAAA | AGCAC | ATTTG | GACCAATTTGGAAGACAAGTACAGATAC 3616 |
| Qy | 1860 | CTTTT | CAAG | CAAGT | GGCAAG | TTTCA | ACAG | AGATTT | TGT | GAC | CGCAGCGTGGGCCCTCCTT 1919 |
| Dd | 3617 | CTTTT | CAAG | CAAGT | GGCAAG | TTTCA | ACAG | AGATTT | TGT | GAC | CGCAGCGTGGGCCCTCCTT 3676 |
| Qy | 1920 | CTGCAT | GATT | CTAT | CCAA | ATTTCCA | AGAC | AGTTGG | GTGA | AGTTGC | ATCCTTTTGGGGCAGT 1979 |
| Dd | 3677 | CTGCAT | GATT | CTAT | CCAA | ATTTCCA | AGAC | AGTTGG | GTGA | AGTTGC | ATCCTTTTGGGGCAGT 3736 |
| Qy | 1980 | AACATT | GAG | CCCA | AGTGT | CCGGA | 2001 | | | | |
| Dd | 3737 | AACATT | GAG | CCCA | AGTGT | CCGGA | 3758 | | | | |

RESULT 3
US-10-149-736-39
; Sequence 39, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.

```
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 5417
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-39

Query Match          91.0%; Score 1820.8; DB 13; Length 5417;
Best Local Similarity 92.6%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 2; Indels 157; Gaps 2;

QY      1  GCGAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTAGAGAA 60
D      1199 GCGAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTAGAGAA 1258
QY      61  GTATTATCGTGGCTTCCTTCTGCTGAGGACACATTCGAAGCACAAAGGAGAGATTTCTAAT 120
D      1259 GTATTATCGTGGCTTCCTTCTGCTGAGGACACATTCGAAGCACAAAGGAGAGATTTCTAAT 1318
QY      121  GATGTGGAAGTGGTGAAGACACAGTTTCTATCTCATGAGGGGTACATGATGGATTGACA 180
D      1319 GATGTGGAAGTGGTGAAGACACAGTTTCTATCTCATGAGGGGTACATGATGGATTGACA 1378
QY      181  GCGCATCAGGCGCGGTTGGTAAATATTCTCAAAATGGGAAGTAAGTGAATGGAAACAGGA 240
D      1379 GCGCATCAGGCGCGGTTGGTAAATATTCTCAAAATGGGAAGTAAGTGAATGGAAACAGGA 1438
QY      241  AAATTATCAGAAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTAAATTCAGA 300
D      1439 AAATTATCAGAAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTAAATTCAGA 1498
QY      301  TGGGAATGCCTCAGGGTAGTGAAGAGCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTA 360
D      1499 TGGGAATGCCTCAGGGTAGTGAAGAGCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTA 1558
QY      361  GATCTCCAGAATC-GAAACTGAAAGAGTTGAATGACTGGCTAACCAAAACAGAGAAAGA 419
D      1559 GATCTCCAGAATCAGAACTGAAAGAGTTGAATGACTGGCTAACCAAAACAGAGAAAGA 1618
QY      420  ACAAGGAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTA 479
D      1619 ACAAGGAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTA 1678
QY      480  CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATCTCTC 539
D      1679 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATCTCTC 1738
QY      540  ACTCACATGGTGGTGGTGAATCTAGTGGAGATCAGGCAACTGCTGCTTTGGAA 599
D      1739 ACTCACATGGTGGTGGTGAATCTAGTGGAGATCAGGCAACTGCTGCTTTGGAA 1798
QY      600  GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCGC 659
D      1799 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCGC 1858
QY      660  TGGGTTCTTTTACAAGACATCTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTT 719
D      1859 TGGGTTCTTTTACAAGACATCTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTT 1918
QY      720  TTTAGTGATGGCTTTTCAGAAAAGAAAGATGCAGTGAACAAGATTTCACACAACTGGCTTT 779
```

```
D      1919 TTTAGTGATGGCTTTTCAGAAAAGAAAGATGCAGTGAACAAGATTTCACACAACCTGGCTTT 1978
QY      780  AAAGATCAAAATGAAATGTTTATCAAGTCTTCAAAAAAACTGGCCGTTTAAAAAGCGGATCTA 839
D      1979 AAAGATCAAAATGAAATGTTTATCAAGTCTTCAAAAAAACTGGCCGTTTAAAAAGCGGATCTA 2038
QY      840  GAAAAGAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACA 899
D      2039 GAAAAGAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACA 2098
QY      900  CTGAAGAATAAAGTCACTGACCCAGAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGT 959
D      2099 CTGAAGAATAAAGTCACTGACCCAGAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGT 2158
QY      960  TGGGATAAATTTAGTCCAAAAAACTTGAAAAGAGTACAGCACAGA----- 1002
D      2159 TGGGATAAATTTAGTCCAAAAAACTTGAAAAGAGTACAGCACAGATTTTCACAGGCTGTCAAC 2218
QY      1003 ----- 1002
D      2219 ACCACTCAGCCATCACTAAACACAGACAACCTGTAATGGAACAGTAACCTACGGTGACCACA 2278
QY      1003 ----- 1002
D      2279 AGGGAACAGATCCTGTAAAGCATGCTCAAGAGGAACTTCCACCACCACCTCCCCAAAAG 2338
QY      1003 -----CCCTTGAAAGACTCCAGGAACTTCAAGAGAGCCACGGATGAG 1043
D      2339 AAGAGGCAGATTACTGTGGATCTTGAAGACTCCAGGAACTTCAAGAGAGCCACGGATGAG 2398
QY      1044  CTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGAT 1103
D      2399 CTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGAT 2458
QY      1104  CTCCTCATTTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGACCTTCGAGGAGAAATT 1163
D      2459 CTCCTCATTTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGACCTTCGAGGAGAAATT 2518
QY      1164  GCGCCTCTGAAAAGAGAACGTGAGCCACGTCATGACCTTGTCTGCCAGCTTACCCTTTG 1223
D      2519 GCGCCTCTGAAAAGAGAACGTGAGCCACGTCATGACCTTGTCTGCCAGCTTACCCTTTG 2578
QY      1224  GGCATTGAGCTCTACCGTATAACCTCAGACTCTCTGGAAGACCTGAAACACAGATGGAAG 1283
D      2579 GGCATTGAGCTCTACCGTATAACCTCAGACTCTCTGGAAGACCTGAAACACAGATGGAAG 2638
QY      1284  CTTCTGCAGGTGGCGTCGAGGACCGAGTCAGGAGCTGATGAGGCCCAAGCCACAGGGACTTT 1343
D      2639 CTTCTGCAGGTGGCGTCGAGGACCGAGTCAGGAGCTGATGAGGCCCAAGCCACAGGGACTTT 2698
QY      1344  GFTCCAGCATCTCAGCACTTTCTTTCCACGTCGTCTCCAGGGTCCCTGGGAGAGAGCCATC 1403
D      2699 GFTCCAGCATCTCAGCACTTTCTTTCCACGTCGTCTCCAGGGTCCCTGGGAGAGAGCCATC 2758
QY      1404  TCGCCAAACAAAGTGCCTTACTATATCAACCACGAGACTCAAAACAACCTTGTGGGACCAT 1463
D      2759 TCGCCAAACAAAGTGCCTTACTATATCAACCACGAGACTCAAAACAACCTTGTGGGACCAT 2818
QY      1464  CCCAAAATGACAGAGCTTACCAGTCTTTTAGCTGACCTGAAATAATGTCAGATTCTCAGCT 1523
D      2819 CCCAAAATGACAGAGCTTACCAGTCTTTTAGCTGACCTGAAATAATGTCAGATTCTCAGCT 2878
QY      1524  TATAGGACTGCCATGAAACTCCGAAGACTGCAGAGAGCCCTTTTGTCTGGATCTTTGAGC 1583
D      2879 TATAGGACTGCCATGAAACTCCGAAGACTGCAGAGAGCCCTTTTGTCTGGATCTTTGAGC 2938
QY      1584  CTGTCAGCTGCATGTGATGCCTTGGACCAGCACAACTCAAGCAAAATGACCAGCCCATG 1643
D      2939 CTGTCAGCTGCATGTGATGCCTTGGACCAGCACAACTCAAGCAAAATGACCAGCCCATG 2998
QY      1644  GATATCCTGCAGATTATTAATTGTTGACCACTATTTATGACCGCCTGGAGCAAGAGCAC 1703
D      2999 GATATCCTGCAGATTATTAATTGTTGACCACTATTTATGACCGCCTGGAGCAAGAGCAC 3058
```


| | | | |
|----|------|--|------|
| QY | 1437 | GAGACTCAAAACAACTTGCTGGGACCATCCAAAATGACAGAGCTCTACCCAGTCTTTAGCT | 1496 |
| Db | 2500 | GAGACTCAAAACAACTTGCTGGGACCATCCAAAATGACAGAGCTCTACCCAGTCTTTAGCT | 2559 |
| QY | 1497 | GACCTGAATAATGTGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAG | 1556 |
| Db | 2560 | GACCTGAATAATGTGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAG | 2619 |
| QY | 1557 | AAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCTTGGACCAGCAC | 1616 |
| Db | 2620 | AAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCTTGGACCAGCAC | 2679 |
| QY | 1617 | AACCTCAAGCAAAAATGACCAGCCCCATGGATATCCTGCAGATTATTAAATTGTTGACCACT | 1676 |
| Db | 2680 | AACCTCAAGCAAAAATGACCAGCCCCATGGATATCCTGCAGATTATTAAATTGTTGACCACT | 2739 |
| QY | 1677 | ATTTATGACCGCCTGGAGCAAGAGCAACAATAATTTGGTCAACGTCCTCTCTCGGTGGAT | 1736 |
| Db | 2740 | ATTTATGACCGCCTGGAGCAAGAGCAACAATAATTTGGTCAACGTCCTCTCTCGGTGGAT | 2799 |
| QY | 1737 | ATGTGTCGAACACTGGCTGCTGAAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTC | 1796 |
| Db | 2800 | ATGTGTCGAACACTGGCTGCTGAAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTC | 2859 |
| QY | 1797 | CTGTCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGTACAGA | 1856 |
| Db | 2860 | CTGTCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGTACAGA | 2919 |
| QY | 1857 | TACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGGCTGGGCCTC | 1916 |
| Db | 2920 | TACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGGCTGGGCCTC | 2979 |
| QY | 1917 | CTTCTGCATGATTTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGC | 1976 |
| Db | 2980 | CTTCTGCATGATTTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGC | 3039 |
| QY | 1977 | AGTAACATTGAGCCAAAGTGTCCGGA | 2001 |
| Db | 3040 | AGTAACATTGAGCCAAAGTGTCCGGA | 3064 |

RESULT 5

```

US-09-845-416-31
; Sequence 31, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 4476
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-31

```

| | | | |
|----|------|--|------|
| Db | 1816 | GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTCTAAT | 1875 |
| Qy | 121 | GATGTGGAAGTGGTGAAGACCAAGTTTCATACTCATGAGGGTACATGATGGATTTTGACA | 180 |
| Db | 1876 | GATGTGGAAGTGGTGAAGACCAAGTTTCATACTCATGAGGGTACATGATGGATTTTGACA | 1935 |
| Qy | 181 | GCCCATCAGGGCCGGTTTGGTAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGA | 240 |
| Db | 1936 | GCCCATCAGGGCCGGTTTGGTAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGA | 1995 |
| Qy | 241 | AAATTATCAGAAGATGAAGAACTGAAGTACAAGACAGATGAATCTCTAAATTCAAGA | 300 |
| Db | 1996 | AAATTATCAGAAGATGAAGAACTGAAGTACAAGACAGATGAATCTCTAAATTCAAGA | 2055 |
| Qy | 301 | TGGGAATGCCTCAGGGTAGCTAGCATGGAATAACAAAGCAATTTACATAGAGTTTTAAATG | 360 |
| Db | 2056 | TGGGAATGCCTCAGGGTAGCTAGCATGGAATAACAAAGCAATTTACATAGAGTTTTAAATG | 2115 |
| Qy | 361 | GATCTCCAGAAATC-GAAACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAA | 419 |
| Db | 2116 | GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAA | 2175 |
| Qy | 420 | ACAAGGAAAAATGGAGGAAGAGCCCTCTTTGGACCTGATCTTTGAAGACCTAAACCGCCAAGTA | 479 |
| Db | 2176 | ACAAGGAAAAATGGAGGAAGAGCCCTCTTTGGACCTGATCTTTGAAGACCTAAACCGCCAAGTA | 2235 |
| Qy | 480 | CAACAACATAAGGTGCTTCAAGAAAGATCTAGAACAAAGAACCAAGTCAGGGTCAATTTCTCTC | 539 |
| Db | 2236 | CAACAACATAAGGTGCTTCAAGAAAGATCTAGAACAAAGAACCAAGTCAGGGTCAATTTCTCTC | 2295 |
| Qy | 540 | ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCCAACTGCTGCTTTGGAA | 599 |
| Db | 2296 | ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCCAACTGCTGCTTTGGAA | 2355 |
| Qy | 600 | GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGACCGC | 659 |
| Db | 2356 | GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGACCGC | 2415 |
| Qy | 660 | TGGGTTCTTTTACAAGACATCCTTCTCAAATGGCAACGTCCTTACTGAAAGAACAGTGCCCTT | 719 |
| Db | 2416 | TGGGTTCTTTTACAAGACAGTTCTTGACCAAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAA | 2475 |
| Qy | 720 | TTTAGTGCATGGCTTTCAGAAAAAGAAAGATGCAGTGAACAAGATTCAACAACCTGGCTTT | 779 |
| Db | 2476 | CTTCTGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGC | 2535 |
| Qy | 780 | AAAGATCAAATGAAATGTTATCAAGTCTTCAAAAACCTGGCCGTTTAAAAAGCGGATCTA | 839 |
| Db | 2536 | GACTTTCAGCAGTTTCAAGAGCAGAACGATGTACATAGGCCCTTCAAGAGGGAATTTGAAA | 2595 |
| Qy | 840 | GAAAAGAAAAAGCAATCCATGGGCAAACTGTA-----TTCACTCAAAACAAG | 885 |
| Db | 2596 | ACTAAGAAACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTCTTGACAGAGCAG | 2655 |
| Qy | 886 | ATCTTCTTTCAACACTGAAGAATAAGTCAGTGACCCAGAAAGACGGAAGCATGGCTGGA-- | 943 |
| Db | 2656 | CCTTTGGAAGGACTAGAGAACTCTACCAGAGCCCAAGAGAGCTGCCTCTCTGAGGAGAGA | 2715 |
| Qy | 944 | -----TAACTTTGCCCGGTGTTGGGATAAATTTAGTCCAAA | 978 |
| Db | 2716 | GCCCAGAAATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATACTAGTGGGAA | 2775 |
| Qy | 979 | AACCTTGAA-----AAGAGTACAGCACAGACCCCTTGAAAGACTC | 1016 |
| Db | 2776 | AAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTC | 2835 |
| Qy | 1017 | CAGGAACCTTCAAGAGCCCAAGGATGAGCTTGACCTCAAGCTGGCCCAAGCTGAGGTGATC | 1076 |
| Db | 2836 | CAGGAACCTTCAAGAGCCCAAGGATGAGCTTGACCTCAAGCTGGCCCAAGCTGAGGTGATC | 2895 |
| Qy | 1077 | AAGGGATCCTGGCAGCCCGTGGCGATCTCTCTCATTTGACTCTCTCTCAAGATCACCTCGAG | 1136 |
| Db | 2896 | AAGGGATCCTGGCAGCCCGTGGCGATCTCTCTCATTTGACTCTCTCTCAAGATCACCTCGAG | 2955 |

| | | | |
|----|------|---|------|
| QY | 885 | GATCTTCTTTCAACACTGAAGAATAAGTCAGTGACCAGAAAGACGGAAAGCATGGCTGGAT | 944 |
| Dd | 1900 | AAAATCCTGAGATCCCTGGAAGGTTCOGATGCAGTCCTGTATTACAAGACGTTTTGGAT | 1959 |
| QY | 945 | AACTTTGCCCGGTGTTGGGATAAATTTAGTCCAAAAAACTTGAAAAGAGTACAGCAC | 999 |
| Dd | 1960 | AACATGAACCTCAAGTGGAGTGAACCTTCGGA AAAAAGTCTCTCAACATTAGTCCCATTG | 2019 |
| QY | 1000 | - - - - - | 999 |
| Dd | 2020 | GAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGG | 2079 |
| QY | 1000 | - - - - - | 999 |
| Dd | 2080 | CTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGCA | 2139 |
| QY | 1000 | - - - - - | 999 |
| Dd | 2140 | GTTCAGAAGCAGAACGATGTACATAGGGCCCTCAAGAGGGAATTGAAAACCTAAAGAACCT | 2199 |
| QY | 1000 | - - - - - | 999 |
| Dd | 2200 | GTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGA | 2259 |
| QY | 1000 | - - - - - | 999 |
| Dd | 2260 | CTAGAGAACTCTACAGGAGCCCCAGAGAGCTGCCTCTTGAGGAGAGGCCAGAAATGTC | 2319 |
| QY | 1000 | - - - - - | 999 |
| Dd | 2320 | ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAATTGAACCTG | 2379 |
| QY | 1000 | - - - - - | 1028 |
| Dd | 2380 | CACCTCCGCTGACTGCGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTCAA | 2439 |
| QY | 1029 | GAGGCCACGGATGAGCTGGACCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGGATCCTGG | 1088 |
| Dd | 2440 | GAGGCCACGGATGAGCTGGACCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGGATCCTGG | 2499 |
| QY | 1089 | CAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCA | 1148 |
| Dd | 2500 | CAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAAGGCA | 2559 |
| QY | 1149 | CTTCGAGGAGAAAATGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGC | 1208 |
| Dd | 2560 | CTTCGAGGAGAAAATGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGC | 2619 |
| QY | 1209 | CAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACCTCTGGAAGACCTG | 1268 |
| Dd | 2620 | CAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACCTCTGGAAGACCTG | 2679 |
| QY | 1269 | AACACCAGATGGAAGCTTCTGTCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAA | 1328 |
| Dd | 2680 | AACACCAGATGGAAGCTTCTGTCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAA | 2739 |
| QY | 1329 | GCCACACAGGGA CTTTGTGTCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCCTC | 1388 |
| Dd | 2740 | GCCACACAGGGA CTTTGTGTCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCCTC | 2799 |
| QY | 1389 | TGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCACGAGACTCAAACA | 1448 |
| Dd | 2800 | TGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCACGAGACTCAAACA | 2859 |
| QY | 1449 | ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGATAAT | 1508 |
| Dd | 2860 | ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGATAAT | 2919 |
| QY | 1509 | GTCAGATTCTCAGCTTATAGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGC | 1568 |
| Dd | 2920 | GTCAGATTCTCAGCTTATAGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGC | 2979 |

RESULT 7

US-09-845-416-29
; Sequence 29, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 4825
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-29

Query Match 63.1%; Score 1263; DB 13; Length 4825;
Best Local Similarity 75.3%; Pred. No. 0;
Matches 1816; Conservative 0; Mismatches 185; Indels 412; Gaps 3;

| | | | |
|----|------|---|------|
| QY | 1 | GGCAGTTCAATTGATGGAGAGTGAAGTAACCTGGACCGTTATCAAACACAGCTTTAGAAGAA | 60 |
| Dd | 1757 | GGCAGTTCAATTGATGGAGAGTGAAGTAACCTGGACCGTTATCAAACACAGCTTTAGAAGAA | 1816 |
| QY | 61 | GTATTATCGTGGCTTCCTTCTGCTGAGGACACATTGCCAAGCACAAAGGAGAGATTTCTTAAT | 120 |
| Dd | 1817 | GTATTATCGTGGCTTCCTTCTGCTGAGGACACATTGCCAAGCACAAAGGAGAGATTTCTTAAT | 1876 |
| QY | 121 | GATGTGGAAGTGCTGAAAAGACCAGTTTTCATACTCATGAGGGGTACATGATGGATTTTGACA | 180 |
| Dd | 1877 | GATGTGGAAGTGCTGAAAAGACCAGTTTTCATACTCATGAGGGGTACATGATGGATTTTGACA | 1936 |
| QY | 181 | GCCCATCAGGGCCGGGTTGGTAATAATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGA | 240 |
| Dd | 1937 | GCCCATCAGGGCCGGGTTGGTAATAATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGA | 1996 |

| | | | |
|----|------|--|------|
| QY | 241 | AAATTATCAGAGATGAAGAACTGAAGTACAAAGACAGATGAATCTCTAAATTCAGA | 300 |
| Db | 1997 | AAATTATCAGAGATGAAGAACTGAAGTACAAAGACAGATGAATCTCTAAATTCAGA | 2056 |
| QY | 301 | TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTACATAGAGTTTAAATG | 360 |
| Db | 2057 | TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTACATAGAGTTTAAATG | 2116 |
| QY | 361 | GATCTCCAGAATC-GAAACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAGAA | 419 |
| Db | 2117 | GATCTCCAGAATCAGAAACTGAAGAGTTGAATGACTGGCTTAACAAAAACAGAAGAA | 2176 |
| QY | 420 | ACAAGGAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAACCGCAAGTA | 479 |
| Db | 2177 | ACAAGGAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAACCGCAAGTA | 2236 |
| QY | 480 | CAACAACATAAGGTGCTTCAAGAAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCCTC | 539 |
| Db | 2237 | CAACAACATAAGGTGCTTCAAGAAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCCTC | 2296 |
| QY | 540 | ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGA | 599 |
| Db | 2297 | ACTCACATGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGA | 2356 |
| QY | 600 | GAACAACCTTAAGGPATTTGGGAGATCGATGGGCAAAACATCTGTAGTGGACAGACCGC | 659 |
| Db | 2357 | GAACAACCTTAAGGPATTTGGGAGATCGATGGGCAAAACATCTGTAGTGGACAGACCGC | 2416 |
| QY | 660 | TGGGTTCTTTTACAAGACATCCTTCTCAATGGCAACGCTCTTACTGAAGAACAGTGCCTT | 719 |
| Db | 2417 | TGGGTTCTTTTACAAGACATCCTATAGATTAAGTCAACAGTCTCCCTGGACCTGGAAAAG | 2476 |
| QY | 720 | TTTAGTGCATGGCTTTTCAAGAAAGAGATGCAGTGAACAAGATTCACACAACCTGGCTTT | 779 |
| Db | 2477 | TTTCTTGCTGGCTTACAGAAGCTGAACAACACTGCCAATGTCTTACAGGATGCTACCCGT | 2536 |
| QY | 780 | A-----AAGATCAAAATGAATGTTATCAAGTCTTCAAAACTGGCCGTT | 824 |
| Db | 2537 | AAGGAAAGGCTCCTAGAAAGACTCCAAGGAGTAAAGAGCTGTATGAACAATGGCAAGAC | 2596 |
| QY | 825 | TTAAAGCGGATCTAGAAAAGAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACA | 884 |
| Db | 2597 | CTCCAAGGTGAATGAAGCTCACACAGATGTTTATCAACAACCTGGATGAACAACAGCCAA | 2656 |
| QY | 885 | GATCTTCTTCAACACTGAAGAAATAAGTCAGTGACCCAGAAAGACGGAAGCATGGCTGGAT | 944 |
| Db | 2657 | AAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCCTGTTACAAAGACGTTTGGAT | 2716 |
| QY | 945 | AACCTTGCCCGGTGTGGGATAATTTAGTCAAAAACTTGAAAAGAGTACAGCAC----- | 999 |
| Db | 2717 | AACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCCATTTG | 2776 |
| QY | 1000 | ----- | 999 |
| Db | 2777 | GAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTCTCTGACGAACTTCTGGTGTGG | 2836 |
| QY | 1000 | ----- | 999 |
| Db | 2837 | CTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGGACTTTCAGCA | 2896 |
| QY | 1000 | ----- | 999 |
| Db | 2897 | GTTCAGAAGCAGAACGATGATACATAGGGCCTTCAAGAGGGAATTGAAAACATAAAGAACCT | 2956 |
| QY | 1000 | ----- | 999 |
| Db | 2957 | GTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGCACAGGACGCTTTTGAAGGA | 3016 |
| QY | 1000 | ----- | 999 |
| Db | 3017 | CTAGAGAAACTCTACCAGAGCCACAGAGCTGCCTCTGTGAGGAGAGAGCCCAAGATGTC | 3076 |

| | | | |
|----|------|---|------|
| QY | 1000 | ----- | 999 |
| Db | 3077 | ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTG | 3136 |
| QY | 1000 | -----AGACCTTTGAAAGACTCCAGGAACCTTCAA | 1028 |
| Db | 3137 | CACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTTCAA | 3196 |
| QY | 1029 | GAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGG | 1088 |
| Db | 3197 | GAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTGG | 3256 |
| QY | 1089 | CAGCCCGTGGCGCATCTCTCATTTGACTCTCTCCAAGATCACTCTGAGAAAGTCAAGGCA | 1148 |
| Db | 3257 | CAGCCCGTGGCGCATCTCTCATTTGACTCTCTCCAAGATCACTCTGAGAAAGTCAAGGCA | 3316 |
| QY | 1149 | CTTCGAGGAGAAATTCGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGC | 1208 |
| Db | 3317 | CTTCGAGGAGAAATTCGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGC | 3376 |
| QY | 1209 | CAGCTTACCCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTG | 1268 |
| Db | 3377 | CAGCTTACCCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTG | 3436 |
| QY | 1269 | AACACCAGATGGAAGCTTCTGCAAGTGGCGCTCGAGGACCGAGTCAGGCAGCTGCATGAA | 1328 |
| Db | 3437 | AACACCAGATGGAAGCTTCTGCAAGTGGCGCTCGAGGACCGAGTCAGGCAGCTGCATGAA | 3496 |
| QY | 1329 | GCCACACAGGACTTTGGTCCAGCATCTCAGCACTTCTTCCACGTGTCCAGGGTCCC | 1388 |
| Db | 3497 | GCCACACAGGACTTTGGTCCAGCATCTCAGCACTTCTTCCACGTGTCCAGGGTCCC | 3556 |
| QY | 1389 | TGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACACGAGACTCAAAACA | 1448 |
| Db | 3557 | TGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACACGAGACTCAAAACA | 3616 |
| QY | 1449 | ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAAT | 1508 |
| Db | 3617 | ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAAT | 3676 |
| QY | 1509 | GTGAGATTCTCAGTTATAGGACTGCCATGAAAACCTCGAAGACTGCAGAAAGGCCCTTTGC | 1568 |
| Db | 3677 | GTGAGATTCTCAGTTATAGGACTGCCATGAAAACCTCGAAGACTGCAGAAAGGCCCTTTGC | 3736 |
| QY | 1569 | TTGGATCTCTTGAGCCTGTGAGTGTGATGCTGATGCTTGTGACCACTATTTATGACCCG | 1628 |
| Db | 3737 | TTGGATCTCTTGAGCCTGTGAGTGTGATGCTGATGCTTGTGACCACTATTTATGACCCG | 3796 |
| QY | 1629 | AATGACCAGCCCATGGATATCCTGCAGATTATTAATTTGTTGACCACTATTTATGACCCG | 1688 |
| Db | 3797 | AATGACCAGCCCATGGATATCCTGCAGATTATTAATTTGTTGACCACTATTTATGACCCG | 3856 |
| QY | 1689 | CTGGAGCAAGAGCACAACAATTTGGTCAACGTCCCTCTCTGCGTGGATATGTGTCTGAAC | 1748 |
| Db | 3857 | CTGGAGCAAGAGCACAACAATTTGGTCAACGTCCCTCTCTGCGTGGATATGTGTCTGAAC | 3916 |
| QY | 1749 | TGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGCTCTCTTTTAAA | 1808 |
| Db | 3917 | TGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGCTCTCTTTTAAA | 3976 |
| QY | 1809 | ACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAG | 1868 |
| Db | 3977 | ACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAG | 4036 |
| QY | 1869 | CAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCTCTCTCTGATGAT | 1928 |
| Db | 4037 | CAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCTCTCTCTGATGAT | 4096 |
| QY | 1929 | TCTATCCAAATTCAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGAGTAACATTGAG | 1988 |
| Db | 4097 | TCTATCCAAATTCAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGAGTAACATTGAG | 4156 |
| QY | 1989 | CCAAGTGTCCGGA | 2001 |

Db 4157 CCAAGTGTCCGGA 4169

RESULT 8
US-09-845-416-35
; Sequence 35, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 4848
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-35

Query Match 63.1%; Score 1263; DB 13; Length 4848;
Best Local Similarity 75.3%; Pred. No. 0;
Matches 1816; Conservative 0; Mismatches 185; Indels 412; Gaps 3;

QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGACCGGTTATCAAAACAGCTTTAGAAAGAA 60
Db 1780 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGACCGGTTATCAAAACAGCTTTAGAAAGAA 1839

QY 61 GTATTATCGTGGCTTCTTCTGCTCAGGACACATTCGAAGCACAAGGAGAGATTTCTAAT 120
Db 1840 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAAT 1899

QY 121 GATGTGGAAGTGGTGAAGACCAAGTTTCATCTCATGAGGGGTACATGATGGATTGACA 180
Db 1900 GATGTGGAAGTGGTGAAGACCAAGTTTCATCTCATGAGGGGTACATGATGGATTGACA 1959

QY 181 GCCCATCAGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGA 240
Db 1960 GCCCATCAGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGA 2019

QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAGACAGATGAATCTCTAAATTCAAGA 300
Db 2020 AAATTATCAGAAGATGAAGAACTGAAGTACAAGACAGATGAATCTCTAAATTCAAGA 2079

QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGATTTTAATG 360
Db 2080 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGATTTTAATG 2139

QY 361 GATCTCCAGAATC-GAAACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAA 419
Db 2140 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAA 2199

QY 420 ACAAGGAAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTA 479
Db 2200 ACAAGGAAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTA 2259

QY 480 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTC 539
Db 2260 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAGGGTCAATTCTCTC 2319

QY 540 ACTCACATGGTGGTGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 599
Db 2320 ACTCACATGGTGGTGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 2379

QY 600 GAACAACCTTAAGGTATTGGGAGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCGC 659
Db 2380 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGC 2439

QY 660 TGGGTTCTTTTACAAGACATCCTTCTCAATGGCAACGTCTTACTGAAGAACAGTGCCTT 719
Db 2440 TGGGTTCTTTTACAAGACATCCTCATAGATTACTGCAACAGTTCCCTCGACCTGGAAAG 2499

QY 720 TTTAGTGCATGGCTTTCAAAAAAGAAAGATGCAAGTGAACAAGATTACACAACACTGGCTTT 779
Db 2500 TTTCTTGCCTGGCTTACAGAAAGCTGAAACAACTGCCAATGTCTTACAGATGCTACCCGT 2559

QY 780 A-----AAGATCAAAATGAATGTTTATCAAGTCTTCAAAAACTGGCCGTT 824
Db 2560 AAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAAGAGCTGATGAACAATGGCAAGAC 2619

QY 825 TTAAGGCGGATCTAGAAAAAGAAAGCAATCCATGGGCAAACTGTATTACTCAAAACA 884
Db 2620 CTCCAAGGTGAAATTAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACAGCAA 2679

QY 885 GATCTTCTTTCAACACTGAAGAAATAAGTCAGTGACCCAGAAAGACGGAGCATGGCTGGAT 944
Db 2680 AAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTGTTACAAAGACGTTTGGAT 2739

QY 945 AACTTTGCCCGGTTGGGATAATTTAGTCCAAAAAACTTGAAAAAGAGTACAGCAC----- 999
Db 2740 AACATGAACCTTCAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAGGTCCCATTTG 2799

QY 1000 ----- 999

Db 2800 GAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTCTCTGCAGGAACCTTCTGGTGTGG 2859

QY 1000 ----- 999

Db 2860 CTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGCA 2919

QY 1000 ----- 999

Db 2920 GTTCAGAAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAAATTGAAAACTAAAGAACCT 2979

QY 1000 ----- 999

Db 2980 GTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGA 3039

QY 1000 ----- 999

Db 3040 CTAGAGAAACTCTACCAGGAGCCAGAGAGTGCCTCCTGAGGAGAGAGCCAGAAATGTC 3099

QY 1000 ----- 999

Db 3100 ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATATCTAGTGGGAAAAATTAACCTG 3159

QY 1000 -----AGACCTTTGAAAGACTCCAGGAACCTTCAA 1028

Db 3160 CACTCCGCTGACTGGCAGAGAAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTTCAA 3219

QY 1029 GAGGCCACGGATGAGCTGGACCTCAAGCTGGCCAAAGCTGAGGTGATCAAGGGATCCTGG 1088

Db 3220 GAGGCCACGGATGAGCTGGACCTCAAGCTGGCCAAAGCTGAGGTGATCAAGGGATCCTGG 3279

QY 1089 CAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCA 1148

Db 3280 CAGCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCA 3339

QY 1149 CTTGAGGAGAAATTGCGCTCTGAAAGAGAACCTGAGCCACCTCAATGACCTTGTCTGC 1208

Db 3340 CTTGAGGAGAAATTGCGCTCTGAAAGAGAACCTGAGCCACCTCAATGACCTTGTCTGC 3399

QY 1209 CAGTTACCACTTTGGGCAATTCAGCTCTCACCCGATTAACCTCAGCACTCTGGAAGACCTG 1268

Db 3400 CAGTTACCACTTTGGGCAATTCAGCTCTCACCCGATTAACCTCAGCACTCTGGAAGACCTG 3459

QY 1269 AACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGTGCATGAA 1328

Db 3460 AACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGTGCATGAA 3519

QY 1329 GCCACAGGGACTTTGTGTCAGCATCTCAGCACTTTCTTCCACGCTCTGTCCAGGGTCCC 1388

Db 3520 GCCACAGGACTTTGGTCCAGCATCTCAGCACTTCTTCCACGTCTGTCCAGGGTCCC 3579
Qy 1389 TGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAACA 1448
Db 3580 TGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAACA 3639
Qy 1449 ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAAT 1508
Db 3640 ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAAT 3699
Qy 1509 GTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGC 1568
Db 3700 GTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGC 3759
Qy 1569 TTGGATCTCTTGAGCCTGTGAGCTGTCATGTGATGCTTGGACCGAGCACAACCTCAAGCAA 1628
Db 3760 TTGGATCTCTTGAGCCTGTGAGCTGTCATGTGATGCTTGGACCGAGCACAACCTCAAGCAA 3819
Qy 1629 AATGACGAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTATGACCCG 1688
Db 3820 AATGACGAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTATGACCCG 3879
Qy 1689 CTGGAGCAAGAGCACACAAATTTGGTCAACGTCCTCTCTGCGTGGATATGTGTCTGAAC 1748
Db 3880 CTGGAGCAAGAGCACACAAATTTGGTCAACGTCCTCTCTGCGTGGATATGTGTCTGAAC 3939
Qy 1749 TGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGTCTTTTAAA 1808
Db 3940 TGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGTCTTTTAAA 3999
Qy 1809 ACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGATACCTTTTCAAG 1868
Db 4000 ACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGATACCTTTTCAAG 4059
Qy 1869 CAAGTGGCAAGTTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCCCTCTTCTGCAATGAT 1928
Db 4060 CAAGTGGCAAGTTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCCCTCTTCTGCAATGAT 4119
Qy 1929 TCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGCAGTAACATTGAG 1988
Db 4120 TCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGCAGTAACATTGAG 4179
Qy 1989 CCAAGTGTCCGGA 2001
Db 4180 CCAAGTGTCCGGA 4192

RESULT 9
US-09-845-416-36
; Sequence 36, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 5060
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-36

Query Match 63.1%; Score 1263; DB 13; Length 5060;
Best Local Similarity 75.3%; Pred. No. 0;
Matches 1816; Conservative 0; Mismatches 185; Indels 412; Gaps 3;

Qy 1 GCAGATTTCATTGATGGAGAGTGAAGTAAACCTTGACCGTTTATCAAACAGCTTTAGAAGAA 60
Db 1992 GCAGATTTCATTGATGGAGAGTGAAGTAAACCTTGACCGTTTATCAAACAGCTTTAGAAGAA 2051
Qy 61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAAT 120
Db 2052 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAAT 2111
Qy 121 GATGTGGAAGTGGTGAAGACCCAGTTTTCATCTACTCATGAGGGGTACATGATGATTTGACA 180
Db 2112 GATGTGGAAGTGGTGAAGACCCAGTTTTCATCTACTCATGAGGGGTACATGATGATTTGACA 2171
Qy 181 GCCCATCAGGCCCGGTTGGTAAATATTTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGA 240
Db 2172 GCCCATCAGGCCCGGTTGGTAAATATTTCTACAATTTGGGAAGTAAGCTGATTGGAACAGGA 2231
Qy 241 AAATTATCAGAAGATGAAGAAAACCTGAAGTACAAGAGCAGATGAATCTCTAAATTTCAAGA 300
Db 2232 AAATTATCAGAAGATGAAGAAAACCTGAAGTACAAGAGCAGATGAATCTCTAAATTTCAAGA 2291
Qy 301 TGGGAATGCCCTCAGGTPAGCTAGCATGGAAAAACAAAAGCAATTTACATAGAGTTTAAATG 360
Db 2292 TGGGAATGCCCTCAGGTPAGCTAGCATGGAAAAACAAAAGCAATTTACATAGAGTTTAAATG 2351
Qy 361 GATCTCCAGAATC-GAAAACCTGAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAGA 419
Db 2352 GATCTCCAGAATCAGAAAACCTGAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAGA 2411
Qy 420 ACAAGGAAAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 479
Db 2412 ACAAGGAAAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 2471
Qy 480 CAACAACATAAAGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAAGTCAATTTCTCTC 539
Db 2472 CAACAACATAAAGTGCTTCAAGAAGATCTAGAACAAAGAACAAAGTCAAGTCAATTTCTCTC 2531
Qy 540 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAAGCAACTGCTGCTTTGGAA 599
Db 2532 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAAGCAACTGCTGCTTTGGAA 2591
Qy 600 GAACAACTTAAGTATTGGGAGATCGATGGGCAAAACATCTGTAGTGAACAGAACCCGC 659
Db 2592 GAACAACTTAAGTATTGGGAGATCGATGGGCAAAACATCTGTAGTGAACAGAACCCGC 2651
Qy 660 TGGGTTCTTTTACAAGACATCCTTCTCAAAATGGCAACGTCTTACTGAAGAACAGTGCCTT 719
Db 2652 TGGGTTCTTTTACAAGACACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAGAAG 2711
Qy 720 TTTAGTGCATGGCTTTCAGAAAAAGAGATGCAGTGAACAAAGATTACACAACTGGCTTT 779
Db 2712 TTTCTTGCCTGGCTTACAGAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCCGT 2771
Qy 780 A-----AAGATCAAAATGAAATGTTTATCAAGTCTTCAAAAAACTGGCCGTT 824
Db 2772 AAGGAAAGGCTCCTAGAAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAAATGGCAAGAC 2831
Qy 825 TTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAA 884
Db 2832 CTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAGCCAA 2891
Qy 885 GATCTTCTTTCAACACTGAAGAATTAAGTCAAGTACCCAGAGACGGAAGCATGGCTGGAT 944
Db 2892 AAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTTACAAAGACGTTTGGAT 2951
Qy 945 AACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAAAATTTGAAAAGAGTACAGCAC----- 999
Db 2952 AACATGAACCTTCAAGTGGAGTGAACTTCGGAAAAAAGTCTCTCAACATTAGGTCCCATTTG 3011
Qy 1000 ----- 999
Db 3012 GAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGTGTGG 3071
Qy 1000 ----- 999

Db 3072 CTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCGGACTTTCCAGCA 3131
QY 1000 ----- 999
Db 3132 GTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCT 3191
QY 1000 ----- 999
Db 3192 GTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAAGGA 3251
QY 1000 ----- 999
Db 3252 CTAGAGAAACTCTACCAGGAGCCAGAGAGTGCCCTCCTGAGGAGAGAGCCAGAAATGTC 3311
QY 1000 ----- 999
Db 3312 ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTG 3371
QY 1000 -----AGACCCCTTGAAAGACTCCAGGAACTTCAA 1028
Db 3372 CACTCCGCTGACTGGCAGAGAAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACCTCAA 3431
QY 1029 GAGGCCACGGATGAGCTGGACCTCAAGCTGGGCAAGCTGAGGTGATCAAGGGATCCTGG 1088
Db 3432 GAGGCCACGGATGAGCTGGACCTCAAGCTGGGCAAGCTGAGGTGATCAAGGGATCCTGG 3491
QY 1089 CAGCCCGTGGCGATCTCCTCATTTGACTCTCTCAAGATCACCTCGAGAAAAAGTCAAGGCA 1148
Db 3492 CAGCCCGTGGCGATCTCCTCATTTGACTCTCTCAAGATCACCTCGAGAAAAAGTCAAGGCA 3551
QY 1149 CTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTGCTCGC 1208
Db 3552 CTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTGCTCGC 3611
QY 1209 CAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATTAACCTCAGCACTCTGGAAGACCTG 1268
Db 3612 CAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATTAACCTCAGCACTCTGGAAGACCTG 3671
QY 1269 AACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGAGCTGCATGAA 1328
Db 3672 AACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGAGCTGCATGAA 3731
QY 1329 GCCACAGGGACTTTGGTCAGCATCTCAGCACTTTCTTCCACGTCGTGCCAGGTCCC 1388
Db 3732 GCCACAGGGACTTTGGTCAGCATCTCAGCACTTTCTTCCACGTCGTGCCAGGTCCC 3791
QY 1389 TGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACACGAGACTCAAACA 1448
Db 3792 TGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACACGAGACTCAAACA 3851
QY 1449 ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTGACCTGAATAAT 1508
Db 3852 ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTGACCTGAATAAT 3911
QY 1509 GTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCAGAGGCCCTTTGTC 1568
Db 3912 GTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCAGAGGCCCTTTGTC 3971
QY 1569 TTGGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCCCTTGGACCAGCACAACTCAAGCAA 1628
Db 3972 TTGGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCCCTTGGACCAGCACAACTCAAGCAA 4031
QY 1629 AATGACCAGCCCATGGATATCCTGCGAGATTATTAATTGTTGACCACCTATTTATGACCCG 1688
Db 4032 AATGACCAGCCCATGGATATCCTGCGAGATTATTAATTGTTGACCACCTATTTATGACCCG 4091
QY 1689 CTGGAGCAAGAGCACAACAAATTTGGTCAACGTCCTCTCTGCGTGGATATGTCTGAAC 1748
Db 4092 CTGGAGCAAGAGCACAACAAATTTGGTCAACGTCCTCTCTGCGTGGATATGTCTGAAC 4151
QY 1749 TGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGCTCTTTTAA 1808

Db 4152 TGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTTTTAA 4211
QY 1809 ACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAG 1868
Db 4212 ACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAG 4271
QY 1869 CAAAGTGGCAAGTTCAAAGGATTTTGTGACCGCAGCGAGGCTGGGCTCCTTCTGCATGAT 1928
Db 4272 CAAAGTGGCAAGTTCAAAGGATTTTGTGACCGCAGCGAGGCTGGGCTCCTTCTGCATGAT 4331
QY 1929 TCTATCCAAATTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGCAGTAACATTGAG 1988
Db 4332 TCTATCCAAATTCCAAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGCAGTAACATTGAG 4391
QY 1989 CCAAGTGTCCGGA 2001
Db 4392 CCAAGTGTCCGGA 4404

RESULT 10
US-10-149-736-40
; Sequence 40, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40
; LENGTH: 5339
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-40

Query Match 62.7%; Score 1254; DB 13; Length 5339;
Best Local Similarity 77.9%; Pred. No. 0;
Matches 1621; Conservative 0; Mismatches 380; Indels 79; Gaps 6;

QY 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA 60
Db 1199 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA 1258
QY 61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACATGAGGAGAGATTTCTAAT 120
Db 1259 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACATGAGGAGAGATTTCTAAT 1318
QY 121 GATGTGGAAGTGGTGAAGACCAAGTTTCATCTCATGAGGGGTACATGATGGATTTGACA 180
Db 1319 GATGTGGAAGTGGTGAAGACCAAGTTTCATCTCATGAGGGGTACATGATGGATTTGACA 1378
QY 181 GCCCATCAGGGCCGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTTGGAACAGGA 240
Db 1379 GCCCATCAGGGCCGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTTGGAACAGGA 1438
QY 241 AAATTATCAGAAAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTTAAATTTCAAGA 300
Db 1439 AAATTATCAGAAAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTTAAATTTCAAGA 1498
QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGACTTTTAATG 360
Db 1499 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATCATAGATTACTG 1558
QY 361 GATCTCCAGAATC-GAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAA 419

Db 1559 CAACAGTTCCCTGGACCTGGAAAAGTTTCTTGGCTGGCTTACAGAAGCTGAAACAACT 1618
QY 420 ACAAGGAAAATGGAGGAGAGCCTCTTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 479
Db 1619 GCCAATGTCCTACAGGATGCTACCCGTAAGGAAGGCTCCTAGAAGACTCCAAGGGAGTA 1678
QY 480 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTCTCTC 539
Db 1679 AAAGAGCTGATGAACAATAGGCAAGACCTCCAGGTGAAATTGAAGCTCACACAGATGTT 1738
QY 540 ACTCACATGGTGGTGTAG-----TTGATGAATCTAGTGGAGATCACGCAACT 587
Db 1739 TATCACAACTGGATGAAAAACAGCCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGAT 1798
QY 588 GCTGCTTTGGAAGAACAACTTAAGG--TATTGGGAGATCGATGGGCAACATCTGTAGA 644
Db 1799 GCAGTCTGTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAGTGAACCTCGGAAA 1858
QY 645 TGGACAGAAGACCGCTGGGTTCTTTTACAAGACATCCTTCTCAAAATGSCAACGTTACT 704
Db 1859 AAGTCTCTCAACATTAGTCCCATTTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCAC 1918
QY 705 GAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGAAAAAAGAAAGATGAGTGAACAAAGATT 764
Db 1919 CTTTCTCTGCAGGAACCTTCTGGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAG 1978
QY 765 CACACAACTGGCTTTTAAAGATCAAAATGAATGTTATCAAGTCTTCAAAAACCTGGCCGTT 824
Db 1979 GCACCTATTGGAGGCGACTTTCAGCAGATTTCAGAAAGCAGAACGATGTACATAGGCGCTT 2038
QY 825 TTAAGCGGATCTAGAAAAAGAAAAAGCAATCCATGGGCAAACTGTA-----871
Db 2039 AAGAGGGAATTGAAAACTAAAGAAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATA 2098
QY 872 -TTCACTCAAAACAGATCTTCTTCAACACTGAAGAAATAAGTCAGTGACCCAGAACGG 930
Db 2099 TTTCTGACAGAGCAGCCTTTTGAAGGACTAGAGAAACTCTACCAGAGCCACAGAGAGCTG 2158
QY 931 AAGCATGGCTGGA-----TAACTTTGCCCGGTGTTGGG 963
Db 2159 CCTCCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTC 2218
QY 964 ATAATTTAGTCCAAAACCTTGAA-----AAGAGTACAGCACAG 1001
Db 2219 AATACTGAGTGGGAAAAAATTGAACCTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAG 2278
QY 1002 ACCCTTGAAAGACTCCAGGAACCTTCAAGAGSCACCGATGAGCTGGAACCTCAAGCTGCGC 1061
Db 2279 ACCCTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGAACCTCAAGCTGCGC 2338
QY 1062 CAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTC 1121
Db 2339 CAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCCTCATTTGACTCTCTC 2398
QY 1122 CAAGATCACCTCGAGAAAAGTCAAGGCACTTCGAGGAGAAATTGGCGCTCTGAAAGAGAAC 1181
Db 2399 CAAGATCACCTCGAGAAAAGTCAAGGCACTTCGAGGAGAAATTGGCGCTCTGAAAGAGAAC 2458
QY 1182 GTGAGCCACGTCAATGACCTTGTCTGCCAGTTTACCCTTTGGGCAATTCAGCTCTCACCG 1241
Db 2459 GTGAGCCACGTCAATGACCTTGTCTGCCAGTTTACCCTTTGGGCAATTCAGCTCTCACCG 2518
QY 1242 TATAACCTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCAAGTGGCCGTC 1301
Db 2519 TATAACCTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCAAGTGGCCGTC 2578
QY 1302 GAGGACCGAGTCAGGAGCTGTCATGAAGCCACAGGGACTTTTGGTCCAGCATCTCAGCAC 1361
Db 2579 GAGGACCGAGTCAGGAGCTGTCATGAAGCCACAGGGACTTTTGGTCCAGCATCTCAGCAC 2638
QY 1362 TTTCTTTCCACGTCTGTCCAGGTTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCC 1421

Db 2639 TTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCC 2698
QY 1422 TACTATATCAACCAAGAGACTCAAACAACTTGTCTGGGACCATCCCAAATGACAGAGCTC 1481
Db 2699 TACTATATCAACCAAGAGACTCAAACAACTTGTCTGGGACCATCCCAAATGACAGAGCTC 2758
QY 1482 TACCAGTCTTTAGCTGACCTGAATAATGTAGATCTCAGATTCTCAGCTTATAGGACTGCCATGAAA 1541
Db 2759 TACCAGTCTTTAGCTGACCTGAATAATGTAGATCTCAGATTCTCAGCTTATAGGACTGCCATGAAA 2818
QY 1542 CTCCGAAGACTGACGAAGGCCCTTTTGTCTTGGATCTCTTGTAGCCTGTGAGCTGATGTGAT 1601
Db 2819 CTCCGAAGACTGACGAAGGCCCTTTTGTCTTGGATCTCTTGTAGCCTGTGAGCTGATGTGAT 2878
QY 1602 GCCTTGGACCAAGCACAACTCAAGCAAAAATGACAGCCCATGGATATCTGAGATTATT 1661
Db 2879 GCCTTGGACCAAGCACAACTCAAGCAAAAATGACAGCCCATGGATATCTGAGATTATT 2938
QY 1662 AATTGTTTGACCACTATTATTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTC 1721
Db 2939 AATTGTTTGACCACTATTATTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTC 2998
QY 1722 CCTCTCTGCGTGGATATGTCTGAACCTGGCTGCTGTAATGTTTATGATACGGGACGAACA 1781
Db 2999 CCTCTCTGCGTGGATATGTCTGAACCTGGCTGCTGTAATGTTTATGATACGGGACGAACA 3058
QY 1782 GGGAGGATCCGTGCTCTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTG 1841
Db 3059 GGGAGGATCCGTGCTCTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTG 3118
QY 1842 GAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAG 1901
Db 3119 GAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAG 3178
QY 1902 CGCAGGCTGGGCTCTCTCTGATGATTTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTT 1961
Db 3179 CGCAGGCTGGGCTCTCTCTGATGATTTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTT 3238
QY 1962 GCATCCTTTGGGGCAGTAACATTTGAGCCAAAGTGTCCGGA 2001
Db 3239 GCATCCTTTGGGGCAGTAACATTTGAGCCAAAGTGTCCGGA 3278

RESULT 11
US-09-845-416-10
; Sequence 10, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 10
; LENGTH: 3531
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-10

Query Match 62.5%; Score 1251.2; DB 13; Length 3531;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1623; Conservative 0; Mismatches 378; Indels 85; Gaps 6;

QY 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTTAGAAGAA 60
Db 1000 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTTAGAAGAA 1059
QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTTGCAAGCACAGGAGAGATTTCCTAAT 120

Db 1060 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAAT 1119
Qy 121 GATGTGGAAGTGGTGAAGACCAAGTTTCATCTCATGAGGGGTACATGATGGATTTGACA 180
Db 1120 GATGTGGAAGTGGTGAAGACCAAGTTTCATCTCATGAGGGGTACATGATGGATTTGACA 1179
Qy 181 GCCCATCAGGCGCGGTTGGTAATATTCTCAAAATGGGAAGTAAGCTGATGGAAACAGGA 240
Db 1180 GCCCATCAGGCGCGGTTGGTAATATTCTCAAAATGGGAAGTAAGCTGATGGAAACAGGA 1239
Qy 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAAGCAGATGAATCTCTTAATTTCAAGA 300
Db 1240 AAATTATCAGAAGATGAAGAACTGAAGTACAAAGCAGATGAATCTCTTAATTTCAAGA 1299
Qy 301 TGGGAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATG 360
Db 1300 TGGGAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAACTCATAGA 1359
Qy 361 GATCTCCAGAA-----TCGAAACTGAAAGAGTTGAATGACTGGCTAACAACAAAACAGAA 413
Db 1360 TTAATGCAACAGTTTCCCTCGACCTGGAAAGTTTCTTGCCTGGCTTACAGAAGCTGAA 1419
Qy 414 GAAAGAACAAAGAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTTGAAGACCTAAACGC 473
Db 1420 ACAACTGCCAATGTCTTACAGGATGCTACCCGTAAAGAAAGGCTCCTAGAAGACTCCAAG 1479
Qy 474 CAACTACAACAACATAAGGTCTTCAAGAAGATCTAGAACAAGAACAAAGTCAGGGTCAAT 533
Db 1480 GGAGTAAAGAGCTGATGAACAATGGCAAGACCTCAAGGTGAATTTGAAGCTCACACA 1539
Qy 534 TCTCTCACTCACATGGTGGTGGTAGTTGATGA-----ATCTAGTGGAGATCAC 581
Db 1540 GATGTTTATCACAACTGGATGAACACAGCCAAATACTCTGAGATCCCTGGAAGTTCC 1599
Qy 582 GCAACTGCTGCTTTGGAAGAACAACTTAAGG---TATTGGGAGATCGATGGGCAACATC 638
Db 1600 GATGATGCAGTCTCTTTACAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAAC 1659
Qy 639 TGATAGTGACAGAAAGACCGCTGGTGGTCTTTTACAGACATCCTTCTCAATGGCAACGT 698
Db 1660 CGGAAAAAGTCTCTCAACATTAGGTCCCATTTTGGAGCCAGTTCTGACCAGTGGAAAGCGT 1719
Qy 699 CTTAATGAAGAACAGTGCCTTTTATAGTGATGCTTTCAGAAAAAGAAAGATCGCAGTGAAC 758
Db 1720 CTGCACTTTCTCTGAGGAATCTGCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 1779
Qy 759 AAGATTACACAACACTGGCTTTTAAAGATCAAAATGAATGTTATCAAGTCTTCAAAAACTG 818
Db 1780 CGGACGACCTATTGGAGGCGACTTTCAGCAGTTTCAAGAGCAGAACGATGTACATAGG 1839
Qy 819 GCCGTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCATGGGCAAACTGTA----- 871
Db 1840 GCCTTCAAGAGGGAAATTGAAACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTA 1899
Qy 872 -----TTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGACCCAGA 924
Db 1900 CGAATATTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGA 1959
Qy 925 AGACGGAAGCATGGCTGGA-----TAACTTTGCCCGGT 957
Db 1960 GAGCTGCCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 2019
Qy 958 GTTGGGATAATTAGTCCAAAAAATTGAA-----AAGAGTACA 995
Db 2020 GAGGTCAATACTAGTGGGAAAAAATTGAACCTGCATCCGCTGACTGGCAGAGAAAAATA 2079
Qy 996 GCACAGACCTTTGAAAGACTCCAGGAATTCAGAGAGCCACGGATGAGCTGGACCTCAAG 1055
Db 2080 GATGAGACCTTTGAAAGACTCCAGGAATTCAGAGAGCCACGGATGAGCTGGACCTCAAG 2139
Qy 1056 CTGGCCAAAGTGAAGTGAATCAAGGATCTTGGCAGCCCGTGGCGGATCTCTCTATTGAC 1115

Db 2140 CTGCGCCAAGCTGAGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTTGAC 2199
Qy 1116 TCTCTCCAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAA 1175
Db 2200 TCTCTCCAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAA 2259
Qy 1176 GAGAACGTGAGCCACGTCAATGACCTTGCTGCCAGCTTACCCTTTGGGCATTCAGCTC 1235
Db 2260 GAGAACGTGAGCCACGTCAATGACCTTGCTGCCAGCTTACCCTTTGGGCATTCAGCTC 2319
Qy 1236 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAAGTGAAGCTTCTGCAGGTG 1295
Db 2320 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAAGTGAAGCTTCTGCAGGTG 2379
Qy 1296 GCGGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTTGGTCCAGCATCT 1355
Db 2380 GCGGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGACTTTTGGTCCAGCATCT 2439
Qy 1356 CAGCACTTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCCCAAAATGACA 1415
Db 2440 CAGCACTTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAA 2499
Qy 1416 GTGCCCTACTATATCAACCACGAGACTCAAAACAACCTTGTGGGACCATCCCAAAATGACA 1475
Db 2500 GTGCCCTACTATATCAACCACGAGACTCAAAACAACCTTGTGGGACCATCCCAAAATGACA 2559
Qy 1476 GAGCTCTACCAGTCTTTAGCTGACCTGAATATGTTCAGATTCTCAGCTTATAGGACTGCC 1535
Db 2560 GAGCTCTACCAGTCTTTAGCTGACCTGAATATGTTCAGATTCTCAGCTTATAGGACTGCC 2619
Qy 1536 ATGAAACTCCGAAGACTGCAGAAAGGCCCTTTTGTGGTGGATCTCTTGAGCCCTGTGAGCTGCA 1595
Db 2620 ATGAAACTCCGAAGACTGCAGAAAGGCCCTTTTGTGGTGGATCTCTTGAGCCCTGTGAGCTGCA 2679
Qy 1596 TGTGATGCCCTTGGACCAAGCAACCTCAAGCAAAATGACCCAGCCCATGGATATCCTGCAG 1655
Db 2680 TGTGATGCCCTTGGACCAAGCAACCTCAAGCAAAATGACCCAGCCCATGGATATCCTGCAG 2739
Qy 1656 ATTATTAATTTGTTGACCACTATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTC 1715
Db 2740 ATTATTAATTTGTTGACCACTATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTC 2799
Qy 1716 AACGTCCCTCTCTGCTGGATATGTGTCTGAATGGCTGCTGAATGTTTATGATACGGGA 1775
Db 2800 AACGTCCCTCTCTGCTGGATATGTGTCTGAATGGCTGCTGAATGTTTATGATACGGGA 2859
Qy 1776 CGAACAGGGAGGATCCGTCTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 1835
Db 2860 CGAACAGGGAGGATCCGTCTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 2919
Qy 1836 CATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGT 1895
Db 2920 CATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGT 2979
Qy 1896 GACCAGCGCAGGCTGGGCTCTCTGCTGATGATTTCTATCCAAATTCACAGAGAGTTGGGT 1955
Db 2980 GACCAGCGCAGGCTGGGCTCTCTGCTGATGATTTCTATCCAAATTCACAGAGAGTTGGGT 3039
Qy 1956 GAAGTTGCATCCTTTTGGGGGAGTAAACATTGAGCCCAAGTGTCCGGA 2001
Db 3040 GAAGTTGCATCCTTTTGGGGGAGTAAACATTGAGCCCAAGTGTCCGGA 3085

RESULT 12
US-09-845-416-30
; Sequence 30, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416

; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 4498
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-30

Query Match 62.5%; Score 1251.2; DB 13; Length 4498;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1623; Conservative 0; Mismatches 378; Indels 85; Gaps 6;

Qy 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAA 60
Db 1757 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAA 1816

Qy 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTAAT 120
Db 1817 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTAAT 1876

Qy 121 GATGTGGAAGTGGTGAAGACCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACA 180
Db 1877 GATGTGGAAGTGGTGAAGACCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACA 1936

Qy 181 GCCCATCAGGGCCGGTGGTAAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240
Db 1937 GCCCATCAGGGCCGGTGGTAAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 1996

Qy 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 300
Db 1997 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 2056

Qy 301 TGGGAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 360
Db 2057 TGGGAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAACTCATAGA 2116

Qy 361 GATCTCCAGAA-----TCGAAACTGAAGAGTTGAATGACTGGCTAAACAAAAACAGAA 413
Db 2117 TTAAGTCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCTGGCTTACAGAAAGCTGAA 2176

Qy 414 GAAAGAAACAAGGAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAACGC 473
Db 2177 ACAACTGCCAATGTCTACAGGATGTACCCGTGAAGAAAGGCTCCTAGAGACTCCAAG 2236

Qy 474 CAAGTACAACAACATAAGGTGCTTCAAGAAGATCTAGAACAGAACAAAGTCAGGGTCAAT 533
Db 2237 GGAGTAAAAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACA 2296

Qy 534 TCTCTCACTCACATGGTGGTGGTAGTTGATGA-----ATCTAGTGGAGATCAC 581
Db 2297 GATGTTTATCACAACTGGATGAAACACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCC 2356

Qy 582 GCAACTGCTGCTTTGGAAGAACAACTTAAG---TATTGGGAGATCGATGGGCAAAACATC 638
Db 2357 GATGATGCAGTCTCTTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAAC TT 2416

Qy 639 TGTAGATGGACAGAAAGCCGCTGTTTATGATGCATGGCTTTTCAAGAAAGAGATGCAGTGAAC 698
Db 2417 CGGAAAAAGTCTCTCAACATTAGTCCCATTTTGGAAAGCCAGTTCTGACCAGTGGAAAGCGT 2476

Qy 699 CTTACTGAAGAACAGTGCCTTTTATGATGCATGGCTTTTCAAGAAAGAGATGCAGTGAAC 758
Db 2477 CTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 2536

Qy 759 AAGATTACACAACCTGGCTTTAAAGATCAAAATGAATGTTATCAAGTCTTCAAAAACCTG 818
Db 2537 CGGCAGGCACCTATTGGAGCGGACTTTCAGCAGCTTTCAGAAAGCAGAACCATGTACATAGG 2596

Qy 819 GCCGTTTAAAAGCGGATCTAGAAAAAGAAAAAGCAATCCATGGGCAAACTGTA----- 871

Db 2597 GCCTTCAAGAGGGAAATTGAAAACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTA 2656
Qy 872 -----TTCACTCAAACAAGATCTTCTTTCAACACACTGAAGAATAAGTCAGTGACCCAGA 924
Db 2657 CGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCAGA 2716
Qy 925 AGACGGAAGCATGGCTGGA-----TAACTTTGCCCGGT 957
Db 2717 GAGTGCCTCTGAGGAGAGAGCCCAAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 2776
Qy 958 GTTGGGATAAATTTAGTCCAAAAAATTGAA-----AAGAGTACA 995
Db 2777 GAGGTCAATACTAGTGGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 2836
Qy 996 GCACAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAG 1055
Db 2837 GATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAG 2896
Qy 1056 CTGCGCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCGCTGGCGATCTCCTCATTGAC 1115
Db 2897 CTGCGCAAAGCTGAGGTGATCAAGGGATCCTGGCAGCCGCTGGCGATCTCCTCATTGAC 2956
Qy 1116 TCTCTCAAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAA 1175
Db 2957 TCTCTCAAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCCTCTGAAA 3016
Qy 1176 GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCCTTTGGGCATTCAAGCTC 1235
Db 3017 GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCCTTTGGGCATTCAAGCTC 3076
Qy 1236 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACACAGATGGAAGCTTCTGCAGGTG 1295
Db 3077 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACACAGATGGAAGCTTCTGCAGGTG 3136
Qy 1296 GCCGTGAGGACCGAGTCAAGGAGCTGATGAAAGCCACAGGGACTTTGGTCCAGCATCT 1355
Db 3137 GCCGTGAGGACCGAGTCAAGGAGCTGATGAAAGCCACAGGGACTTTGGTCCAGCATCT 3196
Qy 1356 CAGCACTTTCTTTCCACGTTCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 1415
Db 3197 CAGCACTTTCTTTCCACGTTCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 3256
Qy 1416 GTGCCCTACTATATCAACCACAGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACA 1475
Db 3257 GTGCCCTACTATATCAACCACAGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACA 3316
Qy 1476 GAGCTTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTTCTCAGCTTATAGGACTGCC 1535
Db 3317 GAGCTTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTTCTCAGCTTATAGGACTGCC 3376
Qy 1536 ATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCA 1595
Db 3377 ATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCA 3436
Qy 1596 TGTGATGCCTTGGACCAAGCACAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAG 1655
Db 3437 TGTGATGCCTTGGACCAAGCACAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAG 3496
Qy 1656 ATTATTAATTGTTTGACCACCTATTATGACCGCTGGAGCAAGAGCACAACAATTTGGTC 1715
Db 3497 ATTATTAATTGTTTGACCACCTATTATGACCGCTGGAGCAAGAGCACAACAATTTGGTC 3556
Qy 1716 AACGTCCTCTCTGCGTGGATATGTTGCTGAACTGGCTGCTGAATGTTTATGATACGGGA 1775
Db 3557 AACGTCCTCTCTGCGTGGATATGTTGCTGAACTGGCTGCTGAATGTTTATGATACGGGA 3616
Qy 1776 CGAACAGGGAGGATCCGTGTCCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 1835
Db 3617 CGAACAGGGAGGATCCGTGTCCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 3676
Qy 1836 CATTGGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGT 1895
Db 3677 CATTGGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGT 3736

| | | | | | |
|----|------|---------------------------|-----------------------|----------------|------|
| QY | 1896 | GACCAGCGCAGGCTGGGCTCTTCTG | CATGATTTCTATCCAAATTC | CAAGACAGTTGGGT | 1955 |
| | | | | | |
| Db | 3737 | GACCAGCGCAGGCTGGGCTCTTCTG | CATGATTTCTATCCAAATTC | CAAGACAGTTGGGT | 3796 |
| | | | | | |
| QY | 1956 | GAAGTTGCATCCTTTGGGGCAGTAA | CATTGAGCCCAAGTCTCCGGA | 2001 | |
| | | | | | |
| Db | 3797 | GAAGTTGCATCCTTTGGGGCAGTAA | CATTGAGCCCAAGTCTCCGGA | 3842 | |
| | | | | | |

QY 1000 -----AGACC 1004
Db 2680 ACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACC 2739
QY 1005 CTTGAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCGCCAA 1064
Db 2740 CTTGAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGCGCCAA 2799
QY 1065 GCTGAGTGATCAAGGGATCCTGGCAGCCCCGTGGCGATCTCCTCATTTGACTCTCTCCAA 1124
Db 2800 GCTGAGTGATCAAGGGATCCTGGCAGCCCCGTGGCGATCTCCTCATTTGACTCTCTCCAA 2859
QY 1125 GATCACTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCCCTCTGAAAGAGAACGTG 1184
Db 2860 GATCACTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCCCTCTGAAAGAGAACGTG 2919
QY 1185 AGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTAT 1244
Db 2920 AGCCACGTCATGACCTTGCTCGCCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTAT 2979
QY 1245 AACCTCAGCACTCTGGAAGACCTGAAACACCAGATGGAAGCTTCTGCAGGTGGCGTCGAG 1304
Db 2980 AACCTCAGCACTCTGGAAGACCTGAAACACCAGATGGAAGCTTCTGCAGGTGGCGTCGAG 3039
QY 1305 GACCGAGTCAGGCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTT 1364
Db 3040 GACCGAGTCAGGCAGCTGCATGAAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTT 3099
QY 1365 CTTTCCACGTCTGTCCAGGTCCTGGGAGAGAGCCATCTCGCCAAACAAGTGCCCTAC 1424
Db 3100 CTTTCCACGTCTGTCCAGGTCCTGGGAGAGAGCCATCTCGCCAAACAAGTGCCCTAC 3159
QY 1425 TATATCAACACCAGAGACTCAAAACAACCTTGCTGGGACCATCCCAAATGACAGAGCTCTAC 1484
Db 3160 TATATCAACACCAGAGACTCAAAACAACCTTGCTGGGACCATCCCAAATGACAGAGCTCTAC 3219
QY 1485 CAGTCTTTAGCTGACCTGAATAATGTGAGATCTCAGCTTATAGGACTGCCATGAAACTC 1544
Db 3220 CAGTCTTTAGCTGACCTGAATAATGTGAGATCTCAGCTTATAGGACTGCCATGAAACTC 3279
QY 1545 CGAAGACTGCAGAAGGCCCTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCC 1604
Db 3280 CGAAGACTGCAGAAGGCCCTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCC 3339
QY 1605 TTGGACCAGCACAACTCAAGCAAAAATGACCCAGCCATGGATATCCTGCAGATTATTAAT 1664
Db 3340 TTGGACCAGCACAACTCAAGCAAAAATGACCCAGCCATGGATATCCTGCAGATTATTAAT 3399
QY 1665 TGTTTGACCACCTATTTATGACCGCCTGGAGCAAGAGCACAACTTTTGTCAACGTCCTT 1724
Db 3400 TGTTTGACCACCTATTTATGACCGCCTGGAGCAAGAGCACAACTTTTGTCAACGTCCTT 3459
QY 1725 CTCTGCGTGGATATGTGTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGG 1784
Db 3460 CTCTGCGTGGATATGTGTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGG 3519
QY 1785 AGGATCCGTGTCCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAA 1844
Db 3520 AGGATCCGTGTCCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAA 3579
QY 1845 GACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGC 1904
Db 3580 GACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGC 3639
QY 1905 AGGCTGGGCCCTCTCTGCATGATTCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCA 1964
Db 3640 AGGCTGGGCCCTCTCTCTGCATGATTCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCA 3699
QY 1965 TCCTTTGGGGCGAGTAACATTGAGCCAAAGTGTCCGGA 2001
Db 3700 TCCTTTGGGGCGAGTAACATTGAGCCAAAGTGTCCGGA 3736

RESULT 14
US-09-845-416-27
; Sequence 27, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 27
; LENGTH: 5149
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-27

Query Match 62.2%; Score 1245; DB 13; Length 5149;
Best Local Similarity 73.1%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 736; Gaps 2;

QY 1 GGCAGTTCAATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAA 60
Db 1757 GGCAGTTCAATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAA 1816
QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTCGAAGCACAAGGAGAGATTCTTAAT 120
Db 1817 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTCGAAGCACAAGGAGAGATTCTTAAT 1876
QY 121 GATGTGGAAGTGGTGAAGACCACTTTCTATCTCATGAGGGGTACATGATGGATTGACA 180
Db 1877 GATGTGGAAGTGGTGAAGACCACTTTCTATCTCATGAGGGGTACATGATGGATTGACA 1936
QY 181 GCCCATCAGGGCCGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240
Db 1937 GCCCATCAGGGCCGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 1996
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAGACAGATGAATCTCTAAATTCAAGA 300
Db 1997 AAATTATCAGAAGATGAAGAACTGAAGTACAAGACAGATGAATCTCTAAATTCAAGA 2056
QY 301 TGGGAATGCCTCAGGTFAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 360
Db 2057 TGGGAATGCCTCAGGTFAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 2116
QY 361 GATCTCCAGAAATC-GAAACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAA 419
Db 2117 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAA 2176
QY 420 ACAAGGAAAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTGAAGACCTAAACCGCAAAGTA 479
Db 2177 ACAAGGAAAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTGAAGACCTAAACCGCAAAGTA 2236
QY 480 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAACAAAGTCAGGGTCAATTCTCTC 539
Db 2237 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAACAAAGTCAGGGTCAATTCTCTC 2296
QY 540 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 599
Db 2297 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 2356
QY 600 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGC 659
Db 2357 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGC 2416
QY 660 TGGGTTCTTTTACAGACATCCTTCTCAATGGCAACGTCCTTACTGAAAGACAGTGCCTT 719
Db 2417 TGGGTTCTTTTACAGACATCCTTCTCAATGGCAACGTCCTTACTGAAAGACAGTGCCTT 2476

QY 720 TTTAGTGCATGGCTTTTCAGAAAAAAGAAAGATGCAGTGAAACAAGATTACACAACTGGCTTT 779
Db |||||
2477 TTTAGTGCATGGCTTTTCAGAAAAAAGAAAGATGCAGTGAAACAAGATTACACAACTGGCTTT 2536
QY 780 AAAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACTGGCCGTTTTTAAAAAGCGGATCTA 839
Db |||||
2537 AAAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACTGGCCGTTTTTAAAAAGCGGATCTA 2596
QY 840 GAAAGAAAAAGCAATCCATGGGCAAACTGTATTCTACTCAAAACAAGATCTTCTTTCAACA 899
Db |||||
2597 GAAAGAAAAAGCAATCCATGGGCAAACTGTATTCTACTCAAAACAAGATCTTCTTTCAACA 2656
QY 900 CTGAAGAATAAGTCAGTGACCCAGAGACGGGAAGCATGGCTGGATAAACTTTTGGCCCGGTGT 959
Db |||||
2657 CTGAAGAATAAGTCAGTGACCCAGAGACGGGAAGCATGGCTGGATAAACTTTTGGCCCGGTGT 2716
QY 960 TGGGATAAATTTAGTCCAAAAAACTTTGAAAAGAGTACAGCAC----- 999
Db |||||
2717 TGGGATAAATTTAGTCCAAAAAACTTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAA 2776
QY 1000 ----- 999
Db CAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGAAAGCTGAAACAACCTGCC 2836
QY 1000 ----- 999
Db AATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAAGACTCCAAGGGAGTAAAA 2896
QY 1000 ----- 999
Db GAGCTGATGAAACAATGGCAAGACCTCCAAAGGTGAAATTGAAGCTCACACAGATGTTTAT 2956
QY 1000 ----- 999
Db CACAACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAGGTTCCGATGATGCA 3016
QY 1000 ----- 999
Db GTCCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAG 3076
QY 1000 ----- 999
Db TCTCTCAACATTAGGTCCCATTTTGGAAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTT 3136
QY 1000 ----- 999
Db TCTCTCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCA 3196
QY 1000 ----- 999
Db CCTATTGGAGGCGACTTTCAGCAGATTTCAGAAAGCAACGATGTACATAGGCGCTTCAAG 3256
QY 1000 ----- 999
Db AGGGAATTGAAACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATT 3316
QY 1000 ----- 999
Db CTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACTCTACCAGGAGGCCAGAGAGCTGCCT 3376
QY 1000 ----- 999
Db CCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAAT 3436
QY 1000 -----AGACC 1004
Db ACTGAGTGGGAAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACC 3496
QY 1005 CTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAA 1064
Db |||||
3497 CTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAA 3556

QY 1065 GCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAA 1124
Db |||||
3557 GCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAA 3616
QY 1125 GATCACCTCGAGAAAAAGTCAAGGCACTTCGAGGAGAAAATTGCGCCTCTGAAAAGAGAACGTG 1184
Db |||||
3617 GATCACCTCGAGAAAAAGTCAAGGCACTTCGAGGAGAAAATTGCGCCTCTGAAAAGAGAACGTG 3676
QY 1185 AGCCACGTCATGACCTTGTCTGCCAGCTTACCACCTTTTGGGCAFTCAGCTCTCACCGTAT 1244
Db |||||
3677 AGCCACGTCATGACCTTGTCTGCCAGCTTACCACCTTTTGGGCAFTCAGCTCTCACCGTAT 3736
QY 1245 AACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAG 1304
Db |||||
3737 AACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAG 3796
QY 1305 GACCGAGTCAGCAGCTGCATGAAGCCACACAGGGACTTTTGGTCCAGCATCTCAGCACTTT 1364
Db |||||
3797 GACCGAGTCAGCAGCTGCATGAAGCCACACAGGGACTTTTGGTCCAGCATCTCAGCACTTT 3856
QY 1365 CTTTCCACGCTGTCTCCAGGGTCCCTGGGAGAGGCCATCTCGCCAAAACAAAGTGCCCTAC 1424
Db |||||
3857 CTTTCCACGCTGTCTCCAGGGTCCCTGGGAGAGGCCATCTCGCCAAAACAAAGTGCCCTAC 3916
QY 1425 TATATCAACCAGAGACTCAAAACAACCTTGTGGACCATCCCAAAATGACAGAGCTCTAC 1484
Db |||||
3917 TATATCAACCAGAGACTCAAAACAACCTTGTGGACCATCCCAAAATGACAGAGCTCTAC 3976
QY 1485 CAGTCTTTAGCTGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTC 1544
Db |||||
3977 CAGTCTTTAGCTGACCTGAATAATGTGAGATTCTCAGCTTATAGGACTGCCATGAAACTC 4036
QY 1545 CGAAGACTGCAGAAAGGCCCTTTTGGTGGATCTCTTGAGCCTGTGAGTGCATGTGATGCC 1604
Db |||||
4037 CGAAGACTGCAGAAAGGCCCTTTTGGTGGATCTCTTGAGCCTGTGAGTGCATGTGATGCC 4096
QY 1605 TTGGACCAGCAACCTCAAGCAAAATGACCCAGCCCATGGATATCCTGCAGATTATTAAAT 1664
Db |||||
4097 TTGGACCAGCAACCTCAAGCAAAATGACCCAGCCCATGGATATCCTGCAGATTATTAAAT 4156
QY 1665 TGTTTGACCACCTATTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCT 1724
Db |||||
4157 TGTTTGACCACCTATTATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCT 4216
QY 1725 CTCTGCGTGGATATGTGTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGG 1784
Db |||||
4217 CTCTGCGTGGATATGTGTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGG 4276
QY 1785 AGGATCCGCTGCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTGGAA 1844
Db |||||
4277 AGGATCCGCTGCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTGGAA 4336
QY 1845 GACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGC 1904
Db |||||
4337 GACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGC 4396
QY 1905 AGGCTGGGCCCTCCTTCTGATGATTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCA 1964
Db |||||
4397 AGGCTGGGCCCTCCTTCTGATGATTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCA 4456
QY 1965 TCCTTTGGGGGCGAGTAACATTGAGCCCAAGTGTCCGGA 2001
Db |||||
4457 TCCTTTGGGGGCGAGTAACATTGAGCCCAAGTGTCCGGA 4493

RESULT 15

US-10-149-736-41
; Sequence 41, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences

FILE REFERENCE: UM-06968
CURRENT APPLICATION NUMBER: US/10/149,736
CURRENT FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: PCT/US01/31126
PRIOR FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: 60/238,848
PRIOR FILING DATE: 2000-10-06
NUMBER OF SEQ ID NOS: 96
SOFTWARE: PatentIn version 3.1
SEQ ID NO 41
LENGTH: 5462
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-10-149-736-41

Query Match 56.6%; Score 1132.2; DB 13; Length 5462;
Best Local Similarity 73.9%; Pred. No. 0;
Matches 1628; Conservative 0; Mismatches 373; Indels 202; Gaps 7;

| | | | |
|----|------|--|------|
| QY | 1 | GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTAGAGAA | 60 |
| Db | 1199 | GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTAGAGAA | 1258 |
| QY | 61 | GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCAAGGAGAGATTTCCTAAT | 120 |
| Db | 1259 | GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCAAGGAGAGATTTCCTAAT | 1318 |
| QY | 121 | GATGTGGAAGTGGTGAAGAACCCAGTTTCATCTCATGAGGGGTACATGATGATTTGACA | 180 |
| Db | 1319 | GATGTGGAAGTGGTGAAGAACCCAGTTTCATCTCATGAGGGGTACATGATGATTTGACA | 1378 |
| QY | 181 | GCCCATCAGGGCCGGTGGTAATATTTCTACAATTTGGAAGTAAGCTGATTGGAACAGGA | 240 |
| Db | 1379 | GCCCATCAGGGCCGGTGGTAATATTTCTACAATTTGGAAGTAAGCTGATTGGAACAGGA | 1438 |
| QY | 241 | AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGA | 300 |
| Db | 1439 | AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGA | 1498 |
| QY | 301 | TGGGAATGCCCTCAGGGTAGCTAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATG | 360 |
| Db | 1499 | TGGGAATGCCCTCAGGGTAGCTAGCATGGAAGAAACAAAGCAATTTACATGCTCCTGGACTG | 1558 |
| QY | 361 | ----- | 360 |
| Db | 1559 | ACCACATATTGGAGCCTCTCCTACTCAGACTGTACTCTGGTGACACAACTGTGGTTACT | 1618 |
| QY | 361 | -----GATCTCCAGAAATCGAAA----- | 377 |
| Db | 1619 | AAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTTCTTGATGTTGGAGCATAGATTA | 1678 |
| QY | 378 | -----CTGAAGAGTTGAATGACTGGCTAAACAAACAAACAGAGAA | 416 |
| Db | 1679 | CTGCAACAGTTCCTCCCTGGACCTGGAAAAGTTTCTTGCTGGCTTACAGAAGCTGAACA | 1738 |
| QY | 417 | AGAACAAAGGAATAAGGAGGAAGAGCCCTTGGACCTGATCTTGAAGACCTAAACGCCAA | 476 |
| Db | 1739 | ACTGCCAATGCTCTACAGGATGCTACCCGTGAAGAAAGGCTCTTAGAAGACTCCAAGGGA | 1798 |
| QY | 477 | GTACAACAATAAAGGTGCTTCAAGAAGATCTAGAACAAAGCAAGTCAAGGTCAATCT | 536 |
| Db | 1799 | GTAAAAGAGCTGATGAACAATAAGGCAAGACCTCCAAGGTGAATTAAGTCAACACAGAT | 1858 |
| QY | 537 | CTCACTCACATGGTGGTGGTAGTTGATGA-----ATCTAGTGGAGATCACGCA | 584 |
| Db | 1859 | GTTTATCACAACCTGGATGAANAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGAT | 1918 |
| QY | 585 | ACTGCTGCTTTGGAAGAACAACTTAAGG--TATTGGGAGATCGATGGCAACATCTGT | 641 |
| Db | 1919 | GATGCAGTCTCTGTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTCGG | 1978 |

| | | | |
|----|------|--|------|
| QY | 542 | AGATGGACAGAAACCGCTGGGTTCTTTTACAAGACATCCTTCTCAAATGGCAACGTCCT | 701 |
| Db | 1979 | AAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTG | 2038 |
| QY | 702 | ACTGAAGAACAGTGCCTTTTATAGTCATGGCTTTTCAGAAAAAGAGATGCAGTGAACAAG | 761 |
| Db | 2039 | CACCTTTCTTCAGAGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGG | 2098 |
| QY | 762 | ATTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTTATCAAGTCTTCAAAAAGTGGCC | 821 |
| Db | 2099 | CAGGCACCTATGGAGGGGACTTTCCAGCAGTTTCAAGCAGAACGATGTACATAGGGCC | 2158 |
| QY | 822 | GTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCATGGGCAAACTGTA----- | 871 |
| Db | 2159 | TTCAAGAGGGAATTGAAAACCTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGA | 2218 |
| QY | 872 | ----TTCACTCAAAACAAGATCTTTCTTTCAACACTGAAGAATAAGTCAGTACCCAGAAGA | 927 |
| Db | 2219 | ATATTTCTGACAGAGCAGCCTTTTGGAGGACTAGAGAACTCTACAGGAGCCCAGAGAG | 2278 |
| QY | 928 | CGGAAGCATGGCTGGA-----TAACTTTGCCCGGTGTT | 960 |
| Db | 2279 | CTGCCTCTGAGGAGAGAGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAG | 2338 |
| QY | 961 | GGGATAAATTTAGTCCAAAAACTTGAA-----AAGAGTACAGCA | 998 |
| Db | 2339 | GTCAATACTGAGTGGGAAAAAATTGAACCTGCACCTCGCTGACTGGCAGAGAAAAATAGAT | 2398 |
| QY | 999 | CAGACCCCTTGAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGTGGACCTCAAGCTG | 1058 |
| Db | 2399 | GAGACCCCTTGAAGACTCCAGGAACCTCAAGAGGCCACGGATGAGTGGACCTCAAGCTG | 2458 |
| QY | 1059 | CGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCT | 1118 |
| Db | 2459 | CGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGGATCTCCTCATTGACTCT | 2518 |
| QY | 1119 | CTCCAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAG | 1178 |
| Db | 2519 | CTCCAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAG | 2578 |
| QY | 1179 | AAGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCA | 1238 |
| Db | 2579 | AAGTGAGCCACGTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCA | 2638 |
| QY | 1239 | CCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAGGTGGCC | 1298 |
| Db | 2639 | CCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAGGTGGCC | 2698 |
| QY | 1299 | GTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGACCTTGGTCCAGCATCTCAG | 1358 |
| Db | 2699 | GTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGACCTTGGTCCAGCATCTCAG | 2758 |
| QY | 1359 | CACCTTTCTTCCACGTCTGTCCAGGTCCTTGGGAGAGAGCCATCTCGCCAAACAAAGTG | 1418 |
| Db | 2759 | CACCTTTCTTCCACGTCTGTCCAGGTCCTTGGGAGAGAGCCATCTCGCCAAACAAAGTG | 2818 |
| QY | 1419 | CCCTACTATATCAACCAACGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAG | 1478 |
| Db | 2819 | CCCTACTATATCAACCAACGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAG | 2878 |
| QY | 1479 | CTCTACCCAGTCTTTAGCTGACCTGAATAATGTGAGATTCCTCAGCTTATAGGACTGCCATG | 1538 |
| Db | 2879 | CTCTACCCAGTCTTTAGCTGACCTGAATAATGTGAGATTCCTCAGCTTATAGGACTGCCATG | 2938 |
| QY | 1539 | AAACTCCGAAGACTGCAGAAGGCCCTTGTGGATCTCTTGGAGCCTGTGAGCTGCATGT | 1598 |
| Db | 2939 | AAACTCCGAAGACTGCAGAAGGCCCTTGTGGATCTCTTGGAGCCTGTGAGCTGCATGT | 2998 |
| QY | 1599 | GATGCCCTTGGACCAAGCAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATT | 1658 |
| Db | 2999 | GATGCCCTTGGACCAAGCAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATT | 3058 |
| QY | 1659 | ATTAATTGTTTGACCACACTATTTATGACCGCCTGGAGCAAGACACAACAAATTTGGTCAAC | 1718 |

| | | | |
|----|------|---|------|
| Db | 3059 | ATTAATTGTTTGACCACTATTATGACCGCCTGGAGCAAGAGCACAAACAATTGGTCAAC | 3118 |
| QY | 1719 | GTCCCTCTCTGCGTGGATATGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGA | 1778 |
| Db | 3119 | GTCCCTCTCTGCGTGGATATGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGA | 3178 |
| QY | 1779 | ACAGGGAGGATCCGTGTCCTGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACAT | 1838 |
| Db | 3179 | ACAGGGAGGATCCGTGTCCTGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACAT | 3238 |
| QY | 1839 | TTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGAC | 1898 |
| Db | 3239 | TTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGAC | 3298 |
| QY | 1899 | CAGCGCAGGCTGGGCCCTCCTTCTGTCATGATTCATCCAAATTCCAAGACAGTTGGGTGAA | 1958 |
| Db | 3299 | CAGCGCAGGCTGGGCCCTCCTTCTGTCATGATTCATCCAAATTCCAAGACAGTTGGGTGAA | 3358 |
| QY | 1959 | GTGTGCATCCTTTGGGGGCAGTAACATTGAGCCCAAGTGTCCGGA | 2001 |
| Db | 3359 | GTGTGCATCCTTTGGGGGCAGTAACATTGAGCCCAAGTGTCCGGA | 3401 |

Search completed: February 2, 2004, 17:40:10
Job time : 697.123 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 11:04:35 ; Search time 3752.09 Seconds
(without alignments)
12961.661 Million cell updates/sec

Title: US-09-845-416-14_COPY_1000_3000
Perfect score: 2001
Sequence: 1 ggcagttcattgatggagag.....cattgagccaagtgtccgga 2001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
- 1: em_estba:*
 - 2: em_esthum:*
 - 3: em_estin:*
 - 4: em_estmu:*
 - 5: em_estov:*
 - 6: em_estpl:*
 - 7: em_estro:*
 - 8: em_htc:*
 - 9: gb_estl:*
 - 10: gb_est2:*
 - 11: gb_htc:*
 - 12: gb_est3:*
 - 13: gb_est4:*
 - 14: gb_est5:*
 - 15: em_estfun:*
 - 16: em_estom:*
 - 17: em_gss_hum:*
 - 18: em_gss_inv:*
 - 19: em_gss_pln:*
 - 20: em_gss_vrt:*
 - 21: em_gss_fun:*
 - 22: em_gss_mam:*
 - 23: em_gss_mus:*
 - 24: em_gss_pro:*
 - 25: em_gss_rod:*
 - 26: em_gss_phg:*
 - 27: em_gss_vrl:*
 - 28: gb_gss1:*
 - 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB ID | Description |
|------------|-------|-------|--------|-------------|--------------------|
| 1 | 991 | 49.5 | 3870 | 11 BC036103 | BC036103 Homo sapi |
| 2 | 776.6 | 38.8 | 3056 | 11 AK044536 | AK044536 Mus muscu |
| 3 | 585.2 | 29.2 | 728 | 14 CB228986 | CB228986 AGENCOURT |
| 4 | 569 | 28.4 | 824 | 9 AL556247 | AL556247 AL556247 |

| | | | | | | |
|----|-------|------|------|----|-----------|--------------------|
| 5 | 567.4 | 28.4 | 1098 | 13 | BX365572 | BX365572 |
| 6 | 566.4 | 28.3 | 801 | 14 | CB991394 | CB991394 AGENCOURT |
| 7 | 558 | 27.9 | 620 | 13 | BQ640063 | BQ640063 he23g04.y |
| 8 | 532.2 | 26.6 | 770 | 10 | BG719710 | BG719710 602690430 |
| 9 | 511.4 | 25.6 | 1298 | 11 | AK087829 | AK087829 Mus muscu |
| 10 | 511.4 | 25.6 | 2135 | 11 | AK013510 | AK013510 Mus muscu |
| 11 | 511.4 | 25.6 | 4437 | 11 | AK036936 | AK036936 Mus muscu |
| 12 | 511 | 25.5 | 797 | 14 | CB960722 | CB960722 AGENCOURT |
| 13 | 509.8 | 25.5 | 1384 | 11 | AK075809 | AK075809 Mus muscu |
| 14 | 494.6 | 24.7 | 652 | 10 | BB629984 | BB629984 BB629984 |
| 15 | 491.6 | 24.6 | 704 | 10 | BB610411 | BB610411 BB610411 |
| 16 | 467.2 | 23.3 | 854 | 9 | AI196693 | AI196693 ui53e10.y |
| 17 | 460.6 | 23.0 | 3753 | 11 | AK081426 | AK081426 Mus muscu |
| 18 | 455.8 | 22.8 | 591 | 2 | HSM083521 | Bx490860 Homo sapi |
| 19 | 455.8 | 22.8 | 717 | 14 | CB527785 | CB527785 UI-M-FY0- |
| 20 | 439 | 21.9 | 578 | 2 | HSM075761 | Bx485574 Homo sapi |
| 21 | 436.6 | 21.8 | 665 | 14 | BY742604 | BY742604 BY742604 |
| 22 | 422.8 | 21.1 | 898 | 10 | BF182065 | BF182065 601804604 |
| 23 | 417.6 | 20.9 | 843 | 14 | CA988247 | CA988247 AGENCOURT |
| 24 | 410.4 | 20.5 | 599 | 10 | BB666688 | BB666688 BB666688 |
| 25 | 407.2 | 20.3 | 554 | 14 | CB613696 | CB613696 AMGNNUC:N |
| 26 | 404.4 | 20.2 | 493 | 14 | CA888041 | CA888041 B0142C06- |
| 27 | 402.6 | 20.1 | 1490 | 11 | BC009242 | BC009242 Homo sapi |
| 28 | 397.4 | 19.9 | 495 | 14 | CA894775 | CA894775 B0187G06- |
| 29 | 397 | 19.8 | 595 | 14 | CB177816 | CB177816 is21c01.x |
| 30 | 392.2 | 19.6 | 650 | 14 | BY714491 | BY714491 BY714491 |
| 31 | 385.2 | 19.3 | 684 | 9 | AL641565 | AL641565 AL641565 |
| 32 | 383.4 | 19.2 | 483 | 14 | CA893902 | CA893902 B0182B01- |
| 33 | 381.8 | 19.1 | 423 | 9 | AA460476 | AA460476 zx61e10.r |
| 34 | 380.4 | 19.0 | 633 | 9 | AL796733 | AL796733 AL796733 |
| 35 | 375.6 | 18.8 | 794 | 13 | BU424348 | BU424348 603234189 |
| 36 | 375.4 | 18.8 | 514 | 9 | AL871560 | AL871560 AL871560 |
| 37 | 365.4 | 18.3 | 655 | 14 | CA558919 | CA558919 K0251G03- |
| 38 | 349.6 | 17.5 | 546 | 4 | BX516597 | Bx516597 RZPD Mus |
| 39 | 345.8 | 17.3 | 3051 | 11 | BC036095 | BC036095 Homo sapi |
| 40 | 337.6 | 16.9 | 541 | 9 | AL894729 | AL894729 AL894729 |
| 41 | 337 | 16.8 | 665 | 10 | BE201973 | BE201973 f103h06.x |
| 42 | 332.6 | 16.6 | 2334 | 11 | BC011062 | BC011062 Mus muscu |
| 43 | 330 | 16.5 | 608 | 9 | AL672616 | AL672616 AL672616 |
| 44 | 329 | 16.4 | 532 | 10 | BE334408 | BE334408 ug97g06.y |
| 45 | 325.6 | 16.3 | 532 | 12 | BJ036615 | BJ036615 BJ036615 |

ALIGNMENTS

| | | | | | | |
|------------|--|-------------|---------|------|--------|-----------------|
| RESULT 1 | BC036103 | BC036103 | 3870 bp | mrna | linear | HTC 04-MAR-2003 |
| LOCUS | Homo sapiens, clone IMAGE:5274415, mRNA. | | | | | |
| DEFINITION | BC036103 | | | | | |
| ACCESSION | BC036103.1 | GI:23271310 | | | | |
| VERSION | HTC | | | | | |
| KEYWORDS | Homo sapiens (human) | | | | | |
| SOURCE | Homo sapiens | | | | | |
| ORGANISM | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | | |
| REFERENCE | 1 (bases 1 to 3870) | | | | | |
| AUTHORS | Strausberg, R. | | | | | |
| TITLE | Direct Submission | | | | | |
| JOURNAL | Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA | | | | | |
| REMARK | NIH-MGC Project URL: http://mgc.nci.nih.gov | | | | | |
| COMMENT | Contact: MGC help desk Email: cgaps-r@mail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN) CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 | | | | | |

Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 48 Row: f Column: 7
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein

This clone has the following problem: retained intron.

FEATURES

source

Location/Qualifiers
1..3870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5274415"
/tissue type="Brain, hippocampus"
/clone_lib="NIH_MGC_95"
/lab_host="DH10B"
/note="Vector: pBluescript"

BASE COUNT 1321 a 677 c 763 g 1109 t

ORIGIN

Query Match 49.5%; Score 991; DB 11; Length 3870;
Best Local Similarity 99.9%; Pred. No. 1.8e-198;
Matches 1002; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGACCGGTTATCAAAACAGCTTTAGAAGAA 60
Db |||||
1143 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGACCGGTTATCAAAACAGCTTTAGAAGAA 1202

QY 61 GTATTATCGTGGCTTCTTCTGTGCTGAGGACACATTCGAAGCACAGGAGAGATTTCTAAT 120
Db |||||
1203 GTATTATCGTGGCTTCTTCTGTGCTGAGGACACATTCGAAGCACAGGAGAGATTTCTAAT 1262

QY 121 GATGTGGAAGTGGTGAAGACCAGTTCATCTCATGAGGGGTACATGATGGATTTGACA 180
Db |||||
1263 GATGTGGAAGTGGTGAAGACCAGTTCATCTCATGAGGGGTACATGATGGATTTGACA 1322

QY 181 GCCCATCAGGCCGGGTGGTAAATATTCTACAATTTGGGAAGTAACTGATTTGAACAGGA 240
Db |||||
1323 GCCCATCAGGCCGGGTGGTAAATATTCTACAATTTGGGAAGTAACTGATTTGAACAGGA 1382

QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAGACGAGATGAATCTCTTAATTCAGA 300
Db |||||
1383 AAATTATCAGAAGATGAAGAACTGAAGTACAAGACGAGATGAATCTCTTAATTCAGA 1442

QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAATG 360
Db |||||
1443 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAATG 1502

QY 361 GATCTCCAGAATC-GAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAA 419
Db |||||
1503 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAA 1562

QY 420 ACAAGGAAAATGGAGGAGAGCCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTA 479
Db |||||
1563 ACAAGGAAAATGGAGGAGAGCCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTA 1622

QY 480 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGCAAGTCAAGGTCAATTTCTCTC 539
Db |||||
1623 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAAGCAAGTCAAGGTCAATTTCTCTC 1682

QY 540 ACTCACATGGTGGTGGTAGTGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 599
Db |||||
1683 ACTCACATGGTGGTGGTAGTGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 1742

QY 600 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAAGACCCG 659
Db |||||
1743 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAAGACCCG 1802

QY 660 TGGGTTCTTTTACAGACATCCTTCTCAATGGCAACGCTTACTGAGAACAGTGCCTT 719
Db |||||
1803 TGGGTTCTTTTACAGACATCCTTCTCAATGGCAACGCTTACTGAGAACAGTGCCTT 1862

QY 720 TTTAGTGCATGGCTTTCAGAAAAAGAGATGCAGTGAACAAGATTTCACAACTGGCTTT 779
Db |||||
1863 TTTAGTGCATGGCTTTCAGAAAAAGAGATGCAGTGAACAAGATTTCACAACTGGCTTT 1922

QY 780 AAAGATCAAAATGAATGTTTATCAAGTCTTCAAAAACTGGCCGTTTAAAAAGCGGATCTA 839
Db |||||
1923 AAAGATCAAAATGAATGTTTATCAAGTCTTCAAAAACTGGCCGTTTAAAAAGCGGATCTA 1982

QY 840 GAAAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAGATCTTTTCAACA 899
Db |||||
1983 GAAAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAGATCTTTTCAACA 2042

QY 900 CTGAAGAATAAGTCAGTCACTGACCCAGAACGGAAGCATGGCTGGATAACTTTGCCCGGTGT 959
Db |||||
2043 CTGAAGAATAAGTCAGTCACTGACCCAGAACGGAAGCATGGCTGGATAACTTTGCCCGGTGT 2102

QY 960 TGGGATAATTTAGTCCAAAAAACTTGAAAAAGAGTACAGCACAGA 1002
Db |||||
2103 TGGGATAATTTAGTCCAAAAAACTTGAAAAAGAGTACAGCACAGA 2145

RESULT 2

AK044536

LOCUS

DEFINITION

AK044536 3056 bp mRNA linear HTC 05-DEC-2002
Mus musculus adult retina cDNA, RIKEN full-length enriched library,
clone:A930019F21 product:dystrophin, muscular dystrophy, full
insert sequence.

ACCESSION

VERSION AK044536.1 GI:26090404

KEYWORDS

SOURCE HTC; CAP trapper.

ORGANISM

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE

AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL

MEDLINE 20499374

PUBMED 11042159

REFERENCE

AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL

MEDLINE 20530913

PUBMED 11076861

REFERENCE

AUTHORS

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,

Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohsuki,S.
and Hayashizaki,Y.

Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)

21085660
11217851

5

The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)
6 (bases 1 to 3056)

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Korino,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gs.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Retina RNA was provided by Dr. Stefano Gustincich (Department of
Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA
02115, USA) whose assistance is gratefully acknowledged. Please
visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.

FEATURES

source
1. .3056
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:A930019F21"
/db_xref="taxon:10090"
/clone="A930019F21"
/tissue_type="retina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"

misc_feature

1. .3056
/note="dystrophin, muscular dystrophy (MGD|MGI:94909,
GB|NM_007868, evidence: BLASTN, 100%, match=999)"

BASE COUNT
ORIGIN

1072 a 567 c 607 g 810 t
Query Match 38.8%; Score 776.6; DB 11; Length 3056;
Best Local Similarity 86.5%; Pred. No. 2.9e-153;

Matches 868; Conservative 0; Mismatches 134; Indels 1; Gaps 1;
QY 1 GGCAGTTCAATTGATGGAGAGTGAAGTAAACCTGGACCCTTATCAAACAGCTTTTAGAAGAA 60
DB 1044 GACAGTTCAATTGATGGAGACGGAAGTAAATCTGGATAGTTACCAAACTGCTTTAGAGAA 1103
QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCAACAAGGAGAGATTCTTAAT 120
DB 1104 GTACTTTTCATGGCTTCTTTCTGCCGAGGATACATTGCGAGCACAAAGGAGAGATTTCAAAT 1163
QY 121 GATGTGAAGTGGTGAAGACCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACA 180
DB 1164 GATGTTGAAGAAAGTGAAGAACAGATTTCATGCTCATGAGGGATTTCATGATGGATCTGACA 1223
QY 181 GCCCATCAGGCCGGGTGGTAAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGA 240
DB 1224 TCTCATCAAGGACTTGTGTGTAATGTTCTACAGTTAGGAAGTCAACTAGTTGGAAAAGGG 1283
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGGA 300
DB 1284 AAATTATCAGAAGATGAAGAACTGAAGTGAAGAACAAATCTCTAAATTCAGGA 1343
QY 301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATG 360
DB 1344 TGGGAATGCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATG 1403
QY 361 GATCTCCAGAATC-GAAACTGAAAGAGTTGAATGACTGGTAAACAAAACAGAAAGAGA 419
DB 1404 GATCTCCAGAATCAGAAATTTAAAGAACTAGATGACTGGTTAACAAAACCTGAAGAGAGA 1463
QY 420 ACAAGGAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 479
DB 1464 ACTAAGAAATGGAGGAAGAGCCCTTGGACCTGATCTTGAAGATCTTAAATGCCAAGTA 1523
QY 480 CAACAACATAAGGTGCTTCAAGAAATCTAGAACAAAGCAACAAAGTCAGGTCAAATCTCTC 539
DB 1524 CAACAACATAAGGTGCTTCAAGAAATCTAGAACAGGAGCAGGTCAAGTCAACTCGCTC 1583
QY 540 ACTCAGATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 599
DB 1584 ACTCAGATGGTGGTAGTTGTTGATGAATCCAGCGGTATCATGCAACAGCTGCTTTGGAA 1643
QY 600 GAACAACCTTAAGGTATTGGGAGATCGATGGCAAAACATCTGTAGATGGACAGAACCCG 659
DB 1644 GAACAACCTTAAGGTACTGGGAGATCGATGGGCAAAATATCTGCAGATGGACTGAAGACC 1703
QY 660 TGGGTTCTTTTACAAGACATCCTTCTCAAATGGCAACGCTTACTGAAGAACAGTGCCTT 719
DB 1704 TGGATTGTTTACAAGATATTCTTCTAAATGGCAGCAATTTACTGAAGAACAGTGCCTT 1763
QY 720 TTTAGTGCATGGCTTTTCAGAAAAAAGAGATGCAGTGAACAAGATTACACAACTGGCTTT 779
DB 1764 TTTAGTACATGGCTTTTCAGAAAAAAGAGATGCAATGAAGAACATTCAGACAAAGTGGCTT 1823
QY 780 AAAGATCAAAATGAAATGTTTATCAAGTCTTCAAATCTTCAAATCTGCGCTTTTAAAGCGGATCTA 839
DB 1824 AAAGATCAAAATGAAATGATGTCAAGTCTTCAAAATATCTACTTTAAATATAGATCTA 1883
QY 840 GAAAGAAAAAGCAATCCATGGGCAAACTGTATTCTCACTCAAAACAAGATCTTCTTTCAACA 899
DB 1884 GAAAGAAAAAGCCCAACCATGGAATAAACTAAGTTCACTCAATCAAGATCTACTTTCCGCA 1943
QY 900 CTGAAGATAAGTCAGTGACCCAGAACGGAAGCATGGCTGGATAACTTTGCCCGGTGT 959
DB 1944 CTGAAGATAAGTCAGTGACTCAAAAGATGGAAATCTGGATGGAAAACTTTGCACACAGT 2003
QY 960 TGGGATAATTTAGTCCAAAAACTTGAAAAAGAGTACAGCACAGA 1002
DB 2004 TGGGACAAATTTAAACCCAAAACTTGAAAAAGAGTTTACAGCACAAA 2046

RESULT 3
CB228986

LOCUS CB228986 728 bp mRNA linear EST 10-FEB-2003
DEFINITION AGENCOURT_1149247 NICHD_Rh_Ov1 Macaca mulatta cDNA clone
IMAGE:6884820 5', mRNA sequence.
ACCESSION CB228986
VERSION CB228986.1 GI:28280564
KEYWORDS EST.
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
Cercopithecinae; Macaca.
REFERENCE 1 (bases 1 to 728)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Eliot Spindel
CDNA Library Preparation: CLONTECH
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM3135 row: e column: 11
High quality sequence stop: 583.
FEATURES
source
1..728
/organism="Macaca mulatta"
/mol_type="mRNA"
/db_xref="taxon:9544"
/clone="IMAGE:6884820"
/tissue_type="Ovary"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD_Rh_Ov1"
/note="Organ: ovary; Vector: pDNR-LIB; Site_1: Sfi I;
Site_2: Sfi I; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.0-4.0 kb. Tissue pooled from
pre-pubertal, post pubertal sn menopausal monkeys.
Constructed by Clontech. Note: this is a NICHD Library."
BASE COUNT 256 a 127 c 162 g 182 t 1 others
ORIGIN
Query Match 29.2%; Score 585.2; DB 14; Length 728;
Best Local Similarity 96.9%; Pred. No. 5.8e-113;
Matches 618; Conservative 0; Mismatches 18; Indels 2; Gaps 2;
QY 195 GGTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATATATCAGAAGA 254
DB 1 GGTGGTAATATTCTACAATTGGGAAGTCAGCTGATTGGAACAGGAAATATCAGAAGA 60
QY 255 TGAAGAACTGAAGTACAAGACAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAG 314
DB 61 TGAAGAACTGAAGTACAAGACAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAG 120
QY 315 GGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATC- 373
DB 121 GGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAATCA 180
QY 374 GAAACTGAAAGAGTTGATGACTGGCTAAACAAAACAGAGAAGAAAGCAAGGAAATGGA 433
DB 181 GAAACTGAAAGAGTTGATGACTGGCTGACAAAACAGAGAAGAAAGCAAGGAAATGGA 240
QY 434 GGAAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACACAAACATAAGGT 493
DB 241 GAAAGAACCCCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACACAAACATAAGGT 300
QY 494 GCTTCAAGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTCTCTCACTCACATGGTGGT 553
DB 301 GCTTCAAGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTCTCTCACTCACATGGTGGT 360
QY 554 GGTAGTTGATGAATCTAGTCGAGATCACGCAACTGCTGCTTTGGGAAGAACAACTTAAGGT 613

Db 361 GGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGACCAACTTAAGGT 420
QY 614 ATTGGGAGATCGATGGCAAAACATCTGTAGATGGACAGAAGACCGTGGTCTTTTACA 673
DB 421 ATTGGGAGATCGATGGCAAAACATCTGCAGATGGACAGAAGACCGTGGTCTTTTACA 480
QY 674 AGACATCCTTCTCAAAATGGCAACGCTTACTACTGAAGAACAGTGCCTTTTAGTCATGGCT 733
DB 481 AGACATCCTTCTCAAAATGGCAACGCTTACTACTGAAGAACAGTGCCTTTTAGTCATGGCT 540
QY 734 TTCAGAAAAAAGAAGATGCAGTGAACAAGATTACACAACT-GGCTTTAAAGATCAAAATG 792
DB 541 TTCAGAAAAAAGAAGATGGTGGTGAACAGGATTACACAACTGGGCTTTAAGGATCAAAATG 600
QY 793 AAATGTTATCAAGTCTTCAAAAAACTGGCGGTTTAAAA 830
DB 601 AAATGTTATCAAGTCTTCAAAAAACTGGCGGTTATGTAA 638
RESULT 4
AL556247 824 bp mRNA linear EST 31-MAY-2003
LOCUS AL556247 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION "cDNA clone CSODK001YB17 5-PRIME, mRNA sequence.
ACCESSION AL556247
VERSION AL556247.2 GI:31278051
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 824)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 15, 2001 this sequence version replaced gi:12898746.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1955.r
Contact : Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CSODK001CA09QP1.
FEATURES
Location/Qualifiers
1..824
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODK001YB17"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/note="1st strand cDNA was primed with a NotI-oligo(dt)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 211 a 210 c 206 g 197 t
ORIGIN
Query Match 28.4%; Score 569; DB 9; Length 824;
Best Local Similarity 100.0%; Pred. No. 1.5e-109;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1433 CCACGAGACTCAAAACAACCTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTT 1492
DB 195 CCACGAGACTCAAAACAACCTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTT 254
QY 1493 AGCTGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 1552
DB 255 AGCTGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 314

```
QY 1553 GCAGAAGGCCCTTGTCTTGGATCTCTTGGACCTGTGTCAGCTGCATGTGATGCCTTGGACCA 1612
|||||
Db 315 GCAGAAGGCCCTTGTCTTGGATCTCTTGGACCTGTGTCAGCTGCATGTGATGCCTTGGACCA 374
|||||
QY 1613 GCACAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTGAC 1672
|||||
Db 375 GCACAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTGAC 434
|||||
QY 1673 CACTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGT 1732
|||||
Db 435 CACTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGT 494
|||||
QY 1733 GGATATGTGCTGAAGTGGCTGCTGTAATGTTATGATACGGGACGAACAGGGAGGATCCG 1792
|||||
Db 495 GGATATGTGCTGAAGTGGCTGCTGTAATGTTATGATACGGGACGAACAGGGAGGATCCG 554
|||||
QY 1793 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA 1852
|||||
Db 555 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA 614
|||||
QY 1853 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACACAGGATTTTGTGACGAGCGCAGGCTGGG 1912
|||||
Db 615 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACACAGGATTTTGTGACGAGCGCAGGCTGGG 674
|||||
QY 1913 CCTCCTTCTGCATGATTTCTATCCAAATTCACACAGATTTGGGTGAAGTTGCATCCTTTGG 1972
|||||
Db 675 CCTCCTTCTGCATGATTTCTATCCAAATTCACACAGATTTGGGTGAAGTTGCATCCTTTGG 734
|||||
QY 1973 GGGCAGTAACATTGAGCCAAAGTGTCGCGA 2001
|||||
Db 735 GGGCAGTAACATTGAGCCAAAGTGTCGCGA 763
|||||
```

```
RESULT 5
BX365572
LOCUS
DEFINITION BX365572 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DK001YB21 5-PRIME, mRNA sequence.
ACCESSION BX365572
VERSION BX365572.1 GI:30366927
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1955.r
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS1AK001ZBIQ1.
Location/Qualifiers
1..1098
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK001YB21"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
```

```
FEATURES
source
286 a 271 c 268 g 262 t 11 others
BASE COUNT
```

```
ORIGIN
Query Match 28.4%; Score 567.4; DB 13; Length 1098;
Best Local Similarity 99.8%; Pred. No. 3.5e-109;
Matches 568; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1433 CCACGAGACTCAAAACAACCTTGTCTGGGACCAATCCCAAAATGACAGAGCTCTACCACTCTTT 1492
|||||
Db 204 CCACGAGACTCAAAACAACCTTGTCTGGGACCAATCCCAAAATGACAGAGCTCTACCACTCTTT 263
|||||
QY 1493 AGCTGACCTGAATATGTGATTTCTCAGTTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACT 1552
|||||
Db 264 AGCTGACCTGAATATGTGATTTCTCAGTTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACT 323
|||||
QY 1553 GCAGAAGGCCCTTGTCTTGGATCTCTTGGACCTGTGATGCTGATGCTTGGACCA 1612
|||||
Db 324 GCAGAAGGCCCTTGTCTTGGATCTCTTGGACCTGTGATGCTGATGCTTGGACCA 383
|||||
QY 1613 GCACAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTGAC 1672
|||||
Db 384 GCACAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTGAC 443
|||||
QY 1673 CACTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGT 1732
|||||
Db 444 CACTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGT 503
|||||
QY 1733 GGATATGTGCTGAAGTGGCTGCTGAATGTTATGATACGGGACGAACAGGGAGGATCCG 1792
|||||
Db 504 GGATATGTGCTGAAGTGGCTGCTGAATGTTATGATACGGGACGAACAGGGAGGATCCG 563
|||||
QY 1793 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA 1852
|||||
Db 564 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA 623
|||||
QY 1853 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACACAGGATTTTGTGACGAGCGCAGGCTGGG 1912
|||||
Db 624 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACACAGGATTTTGTGACGAGCGCAGGCTGGG 683
|||||
QY 1913 CCTCCTTCTGCATGATTTCTATCCAAATTCACACAGATTTGGGTGAAGTTGCATCCTTTGG 1972
|||||
Db 684 CCTCCTTCTGCATGATTTCTATCCAAATTCACACAGATTTGGGTGAAGTTGCATCCTTTGG 743
|||||
QY 1973 GGGCAGTAACATTGAGCCAAAGTGTCGCGA 2001
|||||
Db 744 GGGCAGTAACATTGAGCCAAAGTGTCGCGA 772
|||||
```

```
RESULT 6
CB991394
LOCUS
DEFINITION CB991394 Homo sapiens 801 bp mRNA linear EST 01-MAY-2003
AGENCOURT 13627932 NIH MGC_148 Homo sapiens cDNA clone
IMAGE:30336570 5', mRNA sequence.
ACCESSION CB991394
VERSION CB991394.1 GI:30285818
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
1 (bases 1 to 801)
REFERENCE NIH-MGC
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
```


| | | | |
|---|------|--|------|
| Plate: NDAM360 row: d column: 19 | | Homo sapiens cDNA clone he23g04 5', mRNA sequence. | |
| High quality sequence stop: 621. | | BQ640063 | |
| Location/Qualifiers | | VERSION | |
| 1..801 | | BQ640063.1 | |
| /organism="Homo sapiens" | | GI:21764522 | |
| /mol_type="mRNA" | | EST. | |
| /db_xref="taxon:9606" | | Homo sapiens (human) | |
| /clone="IMAGE:30336570" | | Homo sapiens | |
| /tissue_type="pre-eclamptic placenta" | | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| /lab_host="DH10B Tona" | | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | |
| /clone_lib="NIH_MGC_148" | | 1 (bases 1 to 620) | |
| /note="Organ: placenta; Vector: pBluescriptR; Site 1: all-XhoI; Site 2: BamH; Library is oligo-dT primed and directionally cloned using primer 5'-TTTTTTTTTTTTTNN-3', size-selected for average insert size 2.3 kb and normalized to 10^5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library." | | Wistow, G., Bernstein, S.L., Wyatt, M.K., Ray, S., Behal, A., Touchman, J.W., Bouffard, G., Smith, D. and Peterson, K. | |
| | | Expressed sequence tag analysis of human retina for the NEiBank project: Retbindin, an abundant, novel retinal cDNA and alternative splicing of other retina-preferred gene transcripts | |
| | | Mol. Vis. 8 (4), (2002) In press | |
| | | Contact: Wistow G | |
| | | Section on Molecular Structure and Function | |
| | | National Eye Institute | |
| | | 6/331, NIH, Bethesda, MD 20892-2740, USA | |
| | | Tel: 301 402 3452 | |
| | | Fax: 301 496 0078 | |
| | | Email: graeme@helix.nih.gov | |
| | | Plate: 23 row: g column: 04 | |
| | | Seq primer: M13RP1 reverse primer (ABI). | |
| FEATURES | | Location/Qualifiers | |
| source | | 1..620 | |
| | | /organism="Homo sapiens" | |
| | | /mol_type="mRNA" | |
| | | /db_xref="taxon:9606" | |
| | | /clone="he23g04" | |
| | | /tissue_type="Retina" | |
| | | /dev_stage="Adult" | |
| | | /lab_host="EMDH10B" | |
| | | /clone_lib="Human Retina cDNA (Un-normalized, unamplified)" | |
| | |): hd/he" | |
| | | /note="Organ: Eye; Vector: pSPORT1; Neural retina tissue was dissected from two 80 year old donors with no observed eye disease. 100ug of total RNA was used for library construction. A directionally cloned cDNA library in the pSPORT1 vector (Life technologies) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adaptor [5'-pGACTAGTTCTAGATCGGCGGCCGCT(T)15-3']. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)." | |
| BASE COUNT | | 165 a 148 c 144 g 163 t | |
| ORIGIN | | | |
| Query Match | | 27.9%; Score 558; DB 13; Length 620; | |
| Best Local Similarity | | 100.0%; Pred. No. 3.1e-107; | |
| Matches | | 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | |
| QY | 1444 | AAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTTAGCTGACCTGA | 1503 |
| Db | 1 | AAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTTAGCTGACCTGA | 60 |
| QY | 1504 | ATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAGACTGCAGAAGGCC | 1563 |
| Db | 61 | ATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAGACTGCAGAAGGCC | 120 |
| QY | 1564 | TTTGCTTGGATCTCTTGGAGCCTGTGAGCCTGTGATGCTGATGCTTGGACCAACCTCA | 1623 |
| Db | 121 | TTTGCTTGGATCTCTTGGAGCCTGTGAGCCTGTGATGCTGATGCTTGGACCAACCTCA | 180 |
| QY | 1624 | AGCAAAATGACCAAGCCCATGGATATCCTGAGATTATTAATTTGTTGACCACTATTATG | 1683 |
| Db | 181 | AGCAAAATGACCAAGCCCATGGATATCCTGAGATTATTAATTTGTTGACCACTATTATG | 240 |
| QY | 1684 | ACCGCCTGGAGCAAGCAACAATTTGGTCAACGTCCTCTCTGCGTGGATATGTGTC | 1743 |

| | | | |
|---|------|--|------|
| Plate: NDAM360 row: d column: 19 | | Homo sapiens cDNA clone he23g04 5', mRNA sequence. | |
| High quality sequence stop: 621. | | BQ640063 | |
| Location/Qualifiers | | VERSION | |
| 1..801 | | BQ640063.1 | |
| /organism="Homo sapiens" | | GI:21764522 | |
| /mol_type="mRNA" | | EST. | |
| /db_xref="taxon:9606" | | Homo sapiens (human) | |
| /clone="IMAGE:30336570" | | Homo sapiens | |
| /tissue_type="pre-eclamptic placenta" | | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| /lab_host="DH10B Tona" | | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | |
| /clone_lib="NIH_MGC_148" | | 1 (bases 1 to 620) | |
| /note="Organ: placenta; Vector: pBluescriptR; Site 1: all-XhoI; Site 2: BamH; Library is oligo-dT primed and directionally cloned using primer 5'-TTTTTTTTTTTTTNN-3', size-selected for average insert size 2.3 kb and normalized to 10^5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library." | | Wistow, G., Bernstein, S.L., Wyatt, M.K., Ray, S., Behal, A., Touchman, J.W., Bouffard, G., Smith, D. and Peterson, K. | |
| | | Expressed sequence tag analysis of human retina for the NEiBank project: Retbindin, an abundant, novel retinal cDNA and alternative splicing of other retina-preferred gene transcripts | |
| | | Mol. Vis. 8 (4), (2002) In press | |
| | | Contact: Wistow G | |
| | | Section on Molecular Structure and Function | |
| | | National Eye Institute | |
| | | 6/331, NIH, Bethesda, MD 20892-2740, USA | |
| | | Tel: 301 402 3452 | |
| | | Fax: 301 496 0078 | |
| | | Email: graeme@helix.nih.gov | |
| | | Plate: 23 row: g column: 04 | |
| | | Seq primer: M13RP1 reverse primer (ABI). | |
| FEATURES | | Location/Qualifiers | |
| source | | 1..620 | |
| | | /organism="Homo sapiens" | |
| | | /mol_type="mRNA" | |
| | | /db_xref="taxon:9606" | |
| | | /clone="he23g04" | |
| | | /tissue_type="Retina" | |
| | | /dev_stage="Adult" | |
| | | /lab_host="EMDH10B" | |
| | | /clone_lib="Human Retina cDNA (Un-normalized, unamplified)" | |
| | |): hd/he" | |
| | | /note="Organ: Eye; Vector: pSPORT1; Neural retina tissue was dissected from two 80 year old donors with no observed eye disease. 100ug of total RNA was used for library construction. A directionally cloned cDNA library in the pSPORT1 vector (Life technologies) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adaptor [5'-pGACTAGTTCTAGATCGGCGGCCGCT(T)15-3']. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)." | |
| BASE COUNT | | 204 a 210 c 199 g 188 t | |
| ORIGIN | | | |
| Query Match | | 28.3%; Score 566.4; DB 14; Length 801; | |
| Best Local Similarity | | 99.8%; Pred. No. 5.5e-109; | |
| Matches | | 567; Conservative 0; Mismatches 1; Indels 0; Gaps 0; | |
| QY | 1433 | CCACGAGACTCAAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTT | 1492 |
| Db | 140 | CCACGAGACTCAAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTT | 199 |
| QY | 1493 | AGCTGACCTGAATAATGTCTAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAGACT | 1552 |
| Db | 200 | AGCTGACCTGAATAATGTCTAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAGACT | 259 |
| QY | 1553 | GCAGAAGGCCCTTTGCTTGGATCTCTGAGCCTGTGAGCTGCTGATGCTTGGACCA | 1612 |
| Db | 260 | GCAGAAGGCCCTTTGCTTGGATCTCTGAGCCTGTGAGCTGCTGATGCTTGGACCA | 319 |
| QY | 1613 | GCACAACTCAAGCAAAATGACCCATGGATATCCTGCAATTAATTAATTTGTTGAC | 1672 |
| Db | 320 | GCACAACTCAAGCAAAATGACCCATGGATATCCTGCAATTAATTAATTTGTTGAC | 379 |
| QY | 1673 | CACTATTTATGACCGCTGGAGCAAGACCAACAATTTGGTCAACGTCCTCTCTGCGT | 1732 |
| Db | 380 | CACTATTTATGACCGCTGGAGCAAGACCAACAATTTGGTCAACGTCCTCTCTGCGT | 439 |
| QY | 1733 | GGATATGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAGGAGGATCCG | 1792 |
| Db | 440 | GGATATGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAGGAGGATCCG | 499 |
| QY | 1793 | TGTCCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA | 1852 |
| Db | 500 | TGTCCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA | 559 |
| QY | 1853 | CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGGCTGGG | 1912 |
| Db | 560 | CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGGCTGGG | 619 |
| QY | 1913 | CCTCCTTCTGCATGATCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGG | 1972 |
| Db | 620 | CCTCCTTCTGCATGATCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCCTTTGG | 679 |
| QY | 1973 | GGGCAGTAACATTGAGCCAAAGTGTCCGG | 2000 |
| Db | 680 | GGGCAGTAACATTGAGCCAAAGTGTCCGG | 707 |
| RESULT 7 | | BQ640063 | |
| LOCUS | | he23g04.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he | |
| DEFINITION | | 620 bp mRNA linear EST 15-JUL-2002 | |

Db 241 ACCGCTGGAGCAAGACACAACAATTTGGTCAACGTCCCTCTCTGGCTGGATATGTGTC 300

Qy 1744 TGAACCTGGCTGCTGAATGTTTATGATACGGGACCAACAGGAGGATCCGTGTCTCTGCTTT 1803

Db 301 TGAACCTGGCTGCTGAATGTTTATGATACGGGACCAACAGGAGGATCCGTGTCTCTGCTT 360

Qy 1804 TTAAAACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGACAAGTACAGATACCTTTT 1863

Db 361 TTAAAACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGACAAGTACAGATACCTTTT 420

Qy 1864 TCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCTCCTTCTGC 1923

Db 421 TCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCTCCTTCTGC 480

Qy 1924 ATGATTCTATCCAAATTCOAAGACAGTTGGGTGAAGTTGCATCCCTTTGGGGCGAGTAACA 1983

Db 481 ATGATTCTATCCAAATTCOAAGACAGTTGGGTGAAGTTGCATCCCTTTGGGGCGAGTAACA 540

Qy 1984 TTGAGCCAAAGTGTCGGGA 2001

Db 541 TTGAGCCAAAGTGTCGGGA 558

RESULT 8

BG719710

LOCUS

DEFINITION

602690430F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4822807 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

602690430F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4822807 5', mRNA sequence.

BG719710

GI:13998897

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 770)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLAM10731 row: 1 column: 08

High quality sequence stop: 767.

Location/Qualifiers

1..770

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4822807"

/lab_host="DH10B"

/clone_lib="NIH_MGC_97"

/note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

189 a 199 c 198 g 184 t

FEATURES

source

1..770

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4822807"

/lab_host="DH10B"

/clone_lib="NIH_MGC_97"

/note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

189 a 199 c 198 g 184 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

26.6%; Score 532.2; DB 10; Length 770;

98.1%; Pred. No. 8.9e-102;

Matches 560; Conservative 0; Mismatches 8; Indels 3; Gaps 2;

Qy 1433 CCACGAGACTCAACAACACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTT 1492

Db 150 CCACGAGACTCAACAACACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTT 209

Qy 1493 ACCTGACCTGAATATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 1552

Db 210 AGCTGACCTGAATATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 269

Qy 1553 GCAGAAGGCCCTTTTGTCTGGATCTCTTGTAGCCCTGTCTGATGTCATGTGATGCTTGGACCA 1612

Db 270 GCAGAAGGCCCTTTTGTCTGGATCTCTTGTAGCCCTGTCTGATGTCATGTGATGCTTGGACCA 329

Qy 1613 GCACAA--CCTCAAGCAAAATGACCGCCCATGGATATCCTGCAGATTATTAAATTGTTG 1670

Db 330 GCACAATCCTCAAGTCAGAATGACCGCCCATGGATATCCTGCAGATTATTAAATTGTTG 389

Qy 1671 ACCACTATTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGC 1730

Db 390 ACCACTATTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGC 449

Qy 1731 GTGGATATGTGTGAACCTGGCTGTGAATGTTTATGATACGGGACGAACAGGAGGATC 1790

Db 450 GTGGATATGTGTGAACCTGGCTGTGAATGTTTATGATACGGGACGAACAGGAGGATC 509

Qy 1791 CGTGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAG 1850

Db 510 CGTGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAG 569

Qy 1851 TACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTG 1910

Db 570 TACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGA-TTTGTGACCAGCGCAGGCTG 628

Qy 1911 GGCTCTCTTCTGCATGATTTCTATCCAAATTTCCAGACAGTTGGTGAAATTGTCATCCTTT 1970

Db 629 GGCTCTCTTCTGCATGATTTCTATCCAAATTTCCAGACAGTTGGTGAAATTGTCATCCTTT 688

Qy 1971 GGGGCGAGTAACATTGACCAAGTGTCGGGA 2001

Db 689 GGGGCGAGTAACATAGAGCAAGTGTCGGGA 719

RESULT 9

AK087829

LOCUS

DEFINITION

AK087829

Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN full-length enriched library, clone:E330026B12 product:dystrophin, muscular dystrophy, full insert sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

1298 bp mRNA linear HTC 05-DEC-2002

1

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

2

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

11042159

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,

Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Giessi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G., Tomita,M., Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Aono,H., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barth,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyo-Oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S. and Hayashizaki,Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851
5

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1298)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gs.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, tel:81-45-503-9222, Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.

Location/Qualifiers
1. .1298
/organism="Mus musculus"
/mol_type="mRNA"

/strain="C57BL/6J"
/db_xref="FANTOM_DB:E330026B12"
/db_xref="taxon:10090"
/clone="E330026B12"
/sex="female"
/tissue type="ovary"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev stage="2 days pregnant adult"
120. .1142
/note="unnamed protein product; dystrophin, muscular dystrophy (MGD)|MGI:94909, GB|NM_007868, evidence: BLASTN, 100%, match=999)
putative"
/codon start=1
/protein_id="BAC40019.1"
/db_xref="GI:26352778"
/translation="MREHLKGHETQTTCCWDHPKMTLYQSLADLNNVRFSAVRTAMKL RRLQKALCLDLLSLSAACDLDQHNKQNDQPMIDILQIINCLTIYDRLEQEHNNLVN VPLCDMCLNWLNVYDTGRTGIRVLSFKTGIISLKAHLEDDKYRYLFPQVASSTGF CDQRRLLGLLHDSIQIPRLQGEVASFGGSNIEPSVSCFOPANNKPEIEAALFLDWMR LEPQSMVWLPVLHRVAAETAHQAKNICKECPIIGFYRSLKHFNIDICQSCFFSG RVAKGHKMHYPMVEYCTPTTSGEDVRDFAKVLKNKFKRTKRYFAKHPRMGYLPVQTVLE GDNMET"
BASE COUNT 343 a 332 c 296 g 327 t
ORIGIN

Query Match 25.6%; Score 511.4; DB 11; Length 1298;
Best Local Similarity 93.7%; Pred. No. 2.3e-97;
Matches 533; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 1433 CCACGAGACTCAAAACAACCTTGTGGGACCATCCCAAAATGACAGAGCTTACCAGTCTTT 1492
Db 140 CCACGAGACCCAAACCACTTGTGGGACCCACCCCAAAATGACAGAGCTTACCAGTCTTT 199
QY 1493 AGCTGACCTGAATAATGTCTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 1552
Db 200 AGCTGACCTGAATAATGTCTCAGATTCTCCGGATATAGGACTGCCATGAAAGCTCAGAAGGCT 259
QY 1553 GCAGAAAGCCCTTTTGTGGATCTCTTTGAGCCTGTGAGCTGCTGATGATGCTTGGACCA 1612
Db 260 CCAGAAAGCCCTTTTGTGGATCTCTTTGAGCCTGTGAGCTGCTGATGATGCTTGGACCA 319
QY 1613 GCACAACTCAAGCAAAATGACAGCCCATGGATATCTCTGAGATTATTAATTGTTGAC 1672
Db 320 GCACAACTCAAGCAAAATGACAGCCCATGGATATCTCTGAGATTATTAATTGTTGAC 379
QY 1673 CACTATTTATGACCCGCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTCGGT 1732
Db 380 TACAATTTATGATCGTCTGGAGCAAGAGCACAACAATCTGGTCAATGTCCTCTCTGTGT 439
QY 1733 GGATATGTGTCTGAACCTGGCTGCTGAATGTTATGATACGGGACGACACGGGAGGATCCG 1792
Db 440 GGATATGTGTCTCAACTGGCTTCTCAATGTTTATGATACGGGACGACACGGGAGGATCCG 499
QY 1793 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA 1852
Db 500 TGTCTGTCTTTTAAACTGGCATCATTTCTCTGTGTAAAGCACATTTGGAAGACAAGTA 559
QY 1853 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCAGGCTGGG 1912
Db 560 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACTGGCTTTTGTGACCAAGCAGGCTGGG 619
QY 1913 CCTCTTCTGCATGATCTATCCAAATTCAGACAGTGGGTGAAGTTGATCCTTTGG 1972
Db 620 TCCTCTTCTGCATGATCTATTCAAATCCCAAGACAGTGGGTGAAGTTGCTTCTTCTTGG 679
QY 1973 GGGCAGTAACATTGAGCCAAAGTGTCGGGA 2001
Db 680 GGGCAGTAACATTGAGCCGAGTGTGAGGA 708

RESULT 10
AK013510

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

TITLE
JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL

COMMENT

FEATURES
source

LOCUS AK013510 2135 bp mRNA linear HTC 05-DEC-2002
DEFINITION Mus musculus adult male hippocampus cDNA, RIKEN full-length enriched library, clone:2900010C03 product:dystrophin, muscular dystrophy, full insert sequence.
ACCESSION AK013510
VERSION AK013510.1 GI:12850899
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Onara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
4
AUTHORS Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S. and Hayashizaki,Y.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409 (6821), 685-690 (2001)
MEDLINE 21085660
PUBMED 11217851
5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 2135)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,

Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGAGTTAATTAAATATCCCTCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.
FEATURES
Location/Qualifiers
1. 2135
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM,DB:2900010C03"
/db_xref="MGI:1905833"
/db_xref="taxon:10090"
/clone="2900010C03"
/sex="male"
/tissue type="hippocampus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
misc_feature 1. 2135
/note="dystrophin, muscular dystrophy (MGD|MGI:94909, GB|NM_007868, evidence: BLASTN, 100%, match=999)"
BASE COUNT 563 a 457 c 438 g 677 t
ORIGIN
Query Match 25.6%; Score 511.4; DB 11; Length 2135;
Best Local Similarity 93.7%; Pred. No. 2.5e-97;
Matches 533; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 1433 CCACGAGACTCAACAACACTTGCTGGGACCATCCAAATGACAGAGCTCTACCGTCTTT 1492
Db |||||
Db 144 CCACGAGACCCCAACCACTTGTGTTGGACCAACCCAAATGACAGAGCTCTACCGTCTTT 203
QY 1493 AGCTGACCTGATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 1552
Db |||||
Db 204 AGCTGACCTGATAATGTCAGTTCTCCGGTATAGGACTGCCATGAAAGCTCAGAAGGCT 263
QY 1553 GCAGAAGGCCCTTGTGTTGGATCTCTTGAGCCTGTCAGTGCATGTGATGCCTTGGACCA 1612
Db |||||
Db 264 CCAGAAGGCCCTTGTGTTGGATCTCTTGAGCCTGTCAGTGCATGTGATGCCCTGGACCA 323
QY 1613 GCACAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAAATTGTTGAC 1672
Db |||||
Db 324 GCACAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATAAATTAAGTGTGAC 383
QY 1673 CACTATTATTAGCCGCTTGAGCAAGAGCAACAATTTGGTCAACGCTCCCTCTCTGCGT 1732
Db |||||
Db 384 TACAATTTATGATCGTCTGGAGCAAGAGCAACAATCTGGTCAATGTCCCTCTCTGTGT 443

QY 1733 GGATATGTCGTCGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 1792
|||||
Db 444 GGATATGTCGTCGAACTGGCTTCTCAATGTTTATGATACGGGACGAACAGGGAGGATCCG 503
|||||
QY 1793 TGTCTGTCTCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA 1852
|||||
Db 504 TGTCTGTCTCTTTAAAACTGGCATCATTTCTGTGTAAAGCACACTTGAAGACAAGTA 563
|||||
QY 1853 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACACAGGATTTTGTGACCAGCGCAGGCTGGG 1912
|||||
Db 564 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACACTGGCTTTTGTGACCAGCGTAGGCTGGG 623
|||||
QY 1913 CCTCTTCTGTCATGATTCTATCCAAATTCCAAGACAGTGGTGGTGAAGTTGCATCCTTTGG 1972
|||||
Db 624 TCTTCTTCTGTCATGATTCTATTCAAATCCCAAGACAGTGGTGGTGAAGTTGCTTCTCTTTGG 683
|||||
QY 1973 GGGCAGTAACATTGAGCCCAAGTGTCCGGA 2001
|||||
Db 684 GGGCAGTAACATTGAGCCGAGTGTCCAGGA 712
|||||
RESULT 11
AK036936
LOCUS
DEFINITION
AK036936 4437 bp mRNA linear HTC 05-DEC-2002
Mus musculus adult female vagina cDNA, RIKEN full-length enriched
library, clone:9930028B14 product:dystrophin, muscular dystrophy,
full insert sequence.
ACCESSION
AK036936
VERSION
AK036936.1 GI:26085504
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
PUBMED
10349636
AUTHORS
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
PUBMED
11042159
REFERENCE
3
AUTHORS
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
PUBMED
11076861
REFERENCE
4
AUTHORS
Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G., Tomita,M.,
Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
Baldarelli,R., Barsh,G., Blake,J., Boifelli,D., Bojunga,N.,

Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
and Hayashizaki,Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
PUBMED
11217851
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 4437)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/
FEATURES
Location/Qualifiers
1..4437
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:9930028B14"
/db_xref="taxon:10090"
/clone="9930028B14"
/sex="female"
/tissue_type="vagina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
29..1968
/note="dystrophin, muscular dystrophy (MGD|MGI:94909,
GB|NM_007868, evidence: BLASTN, 100%, match=999)
putative"
4417..4422
/note="putative"
4437
/note="putative"
BASE COUNT 1300 a 932 c 885 g 1320 t
ORIGIN
Query Match 25.6%; Score 511.4; DB 11; Length 4437;
Best Local Similarity 93.7%; Pred. No. 2.7e-97;

| | | | | | | | | | |
|---------|------|----------------|---------------|----------------|--------------|------------------|------|------|----|
| Matches | 533; | Conservative | 0; | Mismatches | 36; | Indels | 0; | Gaps | 0; |
| Qy | 1433 | CCACGAGACTCAAA | CAACTTGTCTGGG | ACCATCCCAAAATG | ACAGAGCTCTAC | CAGTCTTT | 1492 | | |
| Db | 49 | CCACGAGACCCAA | CACTTGTGGGAC | CAACCCCAAAATG | ACAGAGCTCTAC | CAGTCTTT | 108 | | |
| Qy | 1493 | AGCTGACCTGAAT | ATGTCAGATTCTC | AGCTTATAGGACT | GCCATGAAACT | CCGAAGACT | 1552 | | |
| Db | 109 | AGCTGACCTGAAT | ATGTCAGATTCTC | AGCTTATAGGACT | GCCATGAAGCT | CAGAAGGCT | 168 | | |
| Qy | 1553 | GCAGAAGGCCCTT | TGCTGGATCTCTT | GAGCCCTGTCAG | CTGCATGTCCTT | TGGACCA | 1612 | | |
| Db | 169 | CCAGAAGGCCCTT | TGCTGGATCTCTT | GAGCCCTGTCAG | CTGCATGTCCTT | TGGACCA | 228 | | |
| Qy | 1613 | GCACAACCTCAAG | CAAAATGACCA | GCCCCATGGAT | ATCCTGCAGATT | ATTAATTGTTGAC | 1672 | | |
| Db | 229 | GCACAACCTCAAG | CAAAATGACCA | GCCCCATGGAT | ATCCTGCAGATT | ATTAATTGTTGAC | 288 | | |
| Qy | 1673 | CACATATTATGAC | CCGCTGGAGCA | AGAGCACAAATTT | TGGTCAACGTC | CTCTCTGCGT | 1732 | | |
| Db | 289 | TACAATTATGAT | CGTCTGGAGCA | AGAGCACAAATTT | TGGTCAATGTC | CTCTCTGCGT | 348 | | |
| Qy | 1733 | GGATATGTGCTG | AACTGGCTGCTG | AATGTTTATGAT | ACGGGACGAC | GAGGAGGATCCG | 1792 | | |
| Db | 349 | GGATATGTGCTG | AACTGGCTTCTC | AATGTTTATGAT | ACGGGACGAC | GAGGAGGATCCG | 408 | | |
| Qy | 1793 | TGTCCTGTCTTT | TAAACTGGCAT | CATTTCCCTGTG | TAAAGCACATTT | TGGAAGACAAGTA | 1852 | | |
| Db | 409 | TGTCCTGTCTTT | TAAACTGGCAT | CATTTCTGTGT | TAAAGCACATTT | TGGAAGACAAGTA | 468 | | |
| Qy | 1853 | CAGATACCTTTT | CAAGCAAGTGG | CAAGTTCAACAGG | ATTTTGTGAC | CCGCGCAGGCTGGG | 1912 | | |
| Db | 469 | CAGATACCTTTT | CAAGCAAGTGG | CAAGTTCAACTG | GGCTTTTGTG | ACCGAGCTAGGCTGGG | 528 | | |
| Qy | 1913 | CCTCCTTCTGC | ATGATCTATCC | AAATCCACAGAG | TTGGGTGAAGT | TGCATCCTTTGG | 1972 | | |
| Db | 529 | TCTTCTTCTGC | ATGATCTATCC | AAATCCACAGAG | TTGGGTGAAGT | TGCATCCTTTGG | 588 | | |
| Qy | 1973 | GGGCAGTAACAT | TGAGCCAAAGT | GTCCGGA | 2001 | | | | |
| Db | 589 | GGGCAGTAACAT | TGAGCCGAGT | GTTCAGGA | 617 | | | | |

RESULT 12
CB960722
LOCUS
DEFINITION
AGENCOURT 13761995 NIH MGC 147 Homo sapiens cDNA clone
IMAGE:3034479 5', mRNA sequence.
CB960722
CB960722.1 GI:30216839
EST.
SOURCE
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 797)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDAM380 row: n column: 08
High quality sequence stop: 520.
Location/Qualifiers
1. .797

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30344479"
/tissue_type="Human Placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_147"
/note="Organ: placenta; Vector: pBluescriptR; Site_1:
ali-XhoI; Site_2: BamH; Oligo-dT primed using primer
5'-TTTTTTTCTTTTCTTTVN-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NHGRI/NHGRI, National Institutes of Health). Note: This is
a NIH_MGC library."
BASE COUNT 199 a 208 c 199 g 191 t
ORIGIN
Query Match 25.5%; Score 511; DB 14; Length 797;
Best Local Similarity 99.8%; Pred. No. 2.7e-97;
Matches 522; Conservative 0; Mismatches 0; Indels 1;
Gaps 1;
Qy 1433 CCACGAGACTCAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTT 1492
Db 168 CCACGAGACTCAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTT 227
Qy 1493 AGCTGACCTGAATATGTCAGATTCTCAGCTTATAGGACTGCCATGAACCTCCGAAGACT 1552
Db 228 AGCTGACCTGAATATGTCAGATTCTCAGCTTATAGGACTGCCATGAACCTCCGAAGACT 287
Qy 1553 GCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGTGCATGTGATGCCTTGGACCA 1612
Db 288 GCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGTGCATGTGATGCCTTGGACCA 347
Qy 1613 GCACAACCTCAAGCAAAATGACCAAGGACCAAGTGCATGATTAATTAATTTGTTGAC 1672
Db 348 GCACAACCTCAAGCAAAATGACCAAGGACCAAGTGCATGATTAATTAATTTGTTGAC 407
Qy 1673 CACTATTATGACCCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGT 1732
Db 408 CACTATTATGACCCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGT 467
Qy 1733 GGATATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGACAGGAGGATCCG 1792
Db 468 GGATATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGACAGGAGGATCCG 527
Qy 1793 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGAAGACAAGTA 1852
Db 528 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGAAGACAAGTA 587
Qy 1853 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGGCTGGG 1912
Db 588 CAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCGCAGGCTGGG 647
Qy 1913 -CCTCCTTCTGCATGATCTATCCAAATTCGAAGACAGTGGG 1954
Db 648 CCTCCTTCTGCATGATCTATCCAAATTCGAAGACAGTGGG 690
RESULT 13
AK075809
LOCUS
DEFINITION
AK075809 1384 bp mRNA linear HTC 07-DEC-2002
Mus musculus 10 day old male pancreas cDNA, RIKEN full-length
enriched library, clone:1810074E05 product:dystrophin, muscular
dystrophy, full insert sequence.
ACCESSION
AK075809
VERSION
AK075809.1 GI:26344655
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE

AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

REFERENCE

AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159

REFERENCE

AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861

REFERENCE

AUTHORS Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G., Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyooka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S. and Hayashizaki,Y.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409 (6821), 685-690 (2001)
MEDLINE 21085660
PUBMED 11217851

REFERENCE

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 1384)

REFERENCE

AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiracka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome

COMMENT Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.

FEATURES

source Location/Qualifiers
1..1384
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:1810074E05"
/db_xref="MGI:1908112"
/db_xref="taxon:10090"
/clone="1810074E05"
/sex="male"
/tissue type="pancreas"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 day old"
120..1142
/note="unnamed protein product; dystrophin, muscular dystrophy (MGD|MGI:94909, GB|NM_007868, evidence: BLASTN, 100%, match=999)
putative"
/codon_start=1
/protein_id="BAC35977.1"
/db_xref="GI:26344656"
/db_xref="MGI:94909"
/translation="MREHLKGHETQTTCDWHPKMTLYQSLADLNVRFSAYRTAMKL RRLQKALCLDLLSLSAACDALDQHNLKONDQPMDILQINCLTIYDRLEQEHNNLVN VPLQVDMCLNLLNVYDTGRTGRIRVLSFKTIIISLCAHLEDKRYLFKQVASSTGF CDQRLGLLLHDSIQIPRLGEVASFGGSNIEPSVRSFQFANNKPEIEAALFLDWMR LEQSMVWLPVLRVAAAEAKHQAKNICKECPHGFYRSLKHFNYDQCSCFFSG RYAKGHKMHYPMVEYCTPTTSGEDVRDFAKVLKNKFRKRYFAKHPRMGYLPVQTVLE GDNMET"

CDS

BASE COUNT 376 a 354 c 303 g 351 t
ORIGIN
Query Match 25.5%; Score 509.8; DB 11; Length 1384;
Best Local Similarity 93.5%; Pred. No. 5.1e-97;
Matches 532; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
QY 1433 CCACGAGACTCAAACTTGTCTGGGACCATCCCAAATGACAGAGCTTACCAGTCTTT 1492
Db 140 CCACGAGACCCAAACCACTTGTGGGACCAACCCAAATGACAGAGCTTACCAGTCTTT 199
QY 1493 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACT 1552
Db 200 AGCTGACCTGAATAATGTCAGGTTCTCCCGTATAGGACTGCCATGAAGCTCAGAAGGCT 259
QY 1553 GCAGAGGCCCTTGTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCTTGGACCA 1612
Db 260 CCAGAGGCCCTTGTGGATCTCTTGAGCCTGTCCGCTGCATGTGATGCCCTGGACCA 319
QY 1613 GCACAACTCAAGCAAAATGACCAAGCCATGGATATCCTGCAGATTATTATTGTTGAC 1672
Db 320 GCACAACTCAAGCAAAATGACCAAGCCATGGATATCCTGCAGATAATTAAGTGTGAC 379
QY 1673 CACTATTTATGACCGCCTGGAGCAAGAGCACAAATTTGGTCAAGCTCCCTCTCTCGGT 1732
Db 380 TACAATTTATGATCGTCTGGAGCAAGAGCACAAATCTGGTCAATGTCCCTCTCTGTGT 439
QY 1733 GGATATGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 1792
Db 440 GGATATGTCTCAACTGGCTTCTCAATGTTTATGATACGGGACGAACAGGGAGGATCCG 499

Db 585 GGGCA-TAACATTGAGCCGAGTGTACAGGA 612

RESULT 15

BB610411

LOCUS

DEFINITION

musculus cDNA clone 1810074E05 5', mRNA sequence.

ACCESSION

BB610411

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BB610411 704 bp mRNA linear EST 26-OCT-2001
BB610411 RIKEN full-length enriched, 10 day old male pancreas Mus
musculus cDNA clone 1810074E05 5', mRNA sequence.
BB610411
BB610411.1 GI:16451938
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 704)
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630. (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.

FEATURES

source

e mouse tissues.
Location/Qualifiers
1..704
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="1810074E05"
/sex="male"
/tissue_type="pancreas"
/dev_stage="10 day old"
/clone_lib="RIKEN full-length enriched, 10 day old male
pancreas"
173 a 185 c 167 g 179 t

BASE COUNT

ORIGIN

Query Match 24.6%; Score 491.6; DB 10; Length 704;
Best Local Similarity 93.0%; Pred. No. 3.2e-93;
Matches 515; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 1433 CCACGAGACTCAAACTTGTGGGACCACATCCCAAATGACAGAGCTCTACCACTCTTT 1492
Db 137 CCACGAGACCCAAACCACTTGTGGGACCACCCCAAATGACAGAGCTCTACCACTCTTT 196
QY 1493 AGCTGACCTGAATAATGTAGATCTCTGAGCTTATAGGCTGCCATGAAACTCCGAAGACT 1552
Db 197 AGCTGACCTGAATAATGTAGATCTCTGAGCTTATAGGCTGCCATGAAACTCAGAAGGCT 256
QY 1553 GCAGAAAGGCCCTTGTGGATCTCTTGAGCCTGTGAGCTGTGATGCTTGGACCA 1612
Db 257 CCAGAAAGGCCCTTGTGGATCTCTTGAGCCTGTGAGCTGTGATGCTTGGACCA 316
QY 1613 GCACAACTCAAGCAAAATGACCAAGCCCATGGATATCTTGAGATTTAATTGTTGAC 1672
Db 317 GCACAACTCAAGCAAAATGACCAAGCCCATGGATATCTTGAGATTTAATTGTTGAC 376
QY 1673 CACTATTATGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCCTCTCTGCGT 1732
Db 377 TACAATTTATGATCGTCTGGAGCAAGAGCACAACAATCTGCTCAATGTCCTCTCTGTGT 436
QY 1733 GGATATGTCTGAACTGGCTGTGTAATGTTTATGATACGGGACGAACAGGAGGATCCG 1792
Db 437 GGATATGTCTCAACTGGCTGTGTAATGTTTATGATACGGGACGAACAGGAGGATCCG 496
QY 1793 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTTAAAGCACATTTGGAGACAAGTA 1852
Db 497 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTTAAAGCACACTTGGAGACAAGTA 556
QY 1853 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGG 1912
Db 557 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACTGGCTTTTGTGACCAACGTAAGCTGGG 616
QY 1913 CCTCCTTCTGCATGATTTCTATCCAAATTCAGAGACAGTTGGGTGAAGTTGCATCCTTTGG 1972
Db 617 TCTTCTTCTGCATGATTTCTATTCAAATCCCAAGACAGTTGGGTGAAGTTGCATCCTTTGG 676
QY 1973 GGGCAGTAACATTG 1986
Db 677 GGGCAGTAACATTG 690

Search completed: February 2, 2004, 06:03:19
Job time : 3754.42 secs